ÜGK 2016: Assessment of mathematics skills

Documentation of questionnaire-based scales

Stefan Sacchi

Dominique Oesch

TREE, Bern 2017
www.tree.unibe.ch

UNIVERSITÄT BERN

Documentation of questionnaire-based scales

Table of contents

Documentation of questionnaire-based scales

Introduction

The Assessment of the Achievement of Basic Educational Competences (henceforward ÜGK)[1] is a regular monitoring survey designed to assess and analyse competences at the primary and lower secondary level in Switzerland. The assessment surveys are tailored to national educational standards as defined by the HarmoS Agreement (swisseducation.educa.ch/-en/harmos).

The focus of the *ÜGK 2016 survey* is on *mathematical competences at the end of lower secondary education*, i.e., at the end of the 11th year of compulsory schooling.[2] In addition, it also includes a context questionnaire designed to capture student, family and context characteristics relevant to both the acquisition of mathematical competence and students' postcompulsory educational pathways. The latter is important as ÜGK 2016 – being part of the national education monitoring system – also serves as the baseline survey for a national panel study on transitions to postcompulsory education, work and adult life ('Transition from Education to Employment [TREE]' survey, see www.tree.unibe.ch). The student questionnaire is geared to maximising the analytical potential for both objectives.

Given the twofold objective of the ÜGK 2016 questionnaire, a modular design was implemented, with one questionnaire module focusing on characteristics relevant to the acquisition of mathematical competence and a second module tailored to capturing important starting conditions of postcompulsory educational transitions. To limit the time required to take the survey, students alternately received one of two questionnaire versions, each comprising a common core questionnaire with instruments of general interest and a module focussing on one of the two objectives mentioned above (see *section 1* for details).

Both modules include numerous *item-based scales* designed to measure latent (i.e., not directly observable) respondent or context characteristics (see Hascher et al. [2015] for the instrumentation of the maths module and Hupka-Brunner et al. [2015] for the rationales behind the education-monitoring module). The core questionnaire and the overall questionnaire design were jointly specified by both research teams. Instrument selection was almost exclusively confined to instruments validated by previous research in the relevant research fields.

This documentation of scales is intended for data analysts who plan to use the scales included in the ÜGK 2016 questionnaire. For each of these scales, the scientific use file pertaining to ÜGK 2016 includes estimates (i.e., scores) of the individual scale values for all participating students.[3] In this documentation, we describe the procedures and statistical models used to calculate these student scores. The following sections outline relevant aspects of the ÜGK

[1] The official terms in German, French and Italian are:
- Überprüfung des Erreichens der Grundkompetenzen (ÜGK)
- Vérification de l'atteinte des compétences fondamentales (COFO)
- Verifica del raggiungimento delle competenze fondamentali (VECOF)

[2] Including two years of kindergarten or a first learning cycle.

[3] The scientific use file also includes composite variables derived from the questionnaire data, which are documented elsewhere (FORS, 2018).

survey design and database (1), the selection and adaptation of scales (2), the statistical modelling and estimation of student scores (3) and some information on the content and guidelines for the interpretation of the scale-specific results, reliability and quality measures (listed in the tabular appendix starting on page 24) (4). Hurried readers may go directly to the brief practical guidelines for the ÜGK scales on page 20.

1. Survey design and database

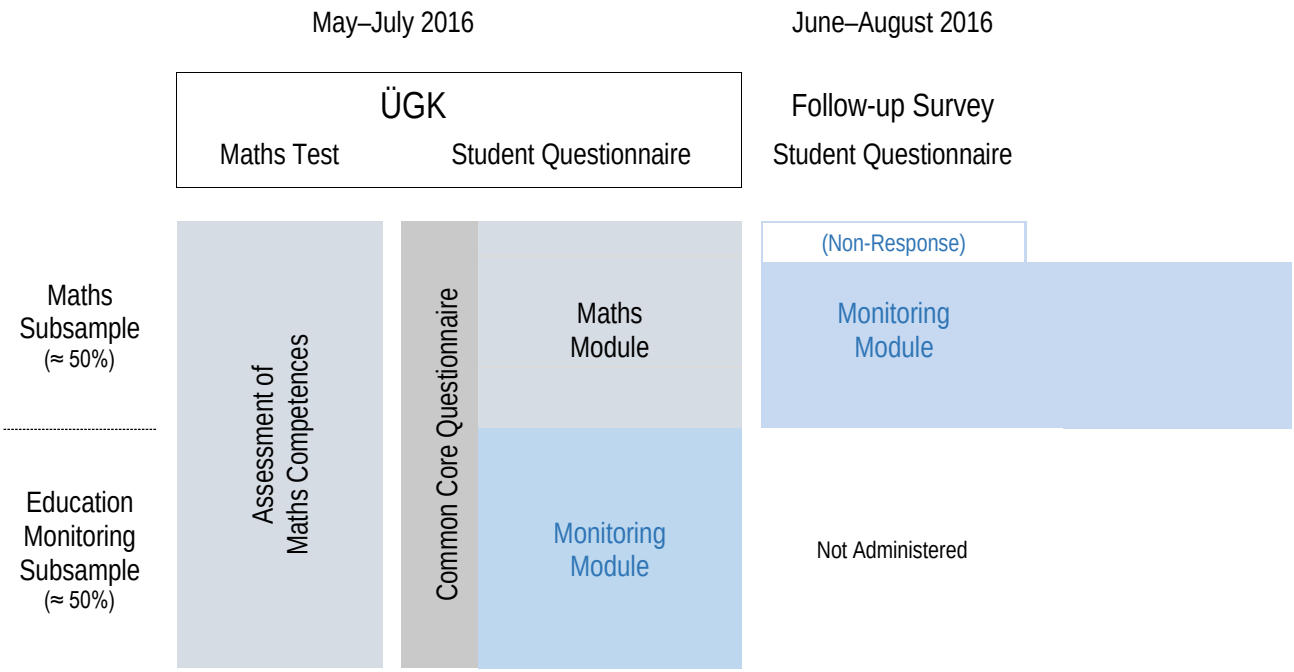
The ÜGK database was collected through a computer-based classroom survey among a random sample of approximately 22,000 students enrolled in the last year of lower secondary education (i.e., the 11th year of compulsory schooling).[4] The survey included two test sessions dedicated to the assessment of basic mathematical competence, along with a computer-assisted self-interview (CASI) of approximately 45 minutes. The student questionnaire covered, among other things, a broad selection of psychometric and other item-based measurement instruments, which is the subject of this documentation.

ÜGK implemented a modular design with two different questionnaire versions, which were carried out among a randomised split-half of the total sample each.[5] The central building block of one questionnaire version is the *'maths module'*, which mainly covered student, teacher and classroom characteristics relevant to the successful acquisition of mathematical competences during compulsory education and to related didactical and pedagogical research. The core of the second questionnaire version is the *'education monitoring module'* designed to collect information on a broad range of resources of the surveyed students, their families and the schools they were attending at the moment of the survey, as these resources are prerequisites for the students' transitions after the completion of compulsory education. The latter module is part of the Swiss Education Monitoring and serves as the baseline survey for the *Transition from Education to Employment-Panel (TREE),* a longitudinal study on post-compulsory pathways of (compulsory) school-leavers (see www.tree.unibe.ch). Both questionnaire versions include a common core section that was answered by all students participating in ÜGK. It incorporates questionnaire items that are relevant to both of the aforementioned research aims pursued with the ÜGK questionnaire.

[4] See Verner and Helbling (forthcoming) for a detailed description of the population.

[5] The random assignment of the students to one questionnaire version is to guarantee that – within each school and each test session – both versions are evenly distributed over the 13 different test booklets used for the preceding mathematics assessment. Hence, from the students' perspective, booklet and questionnaire version are two independent, fully exogenous conditions.

Figure 1 Design and follow-up survey of ÜGK 2016



A follow-up survey among those students who received the questionnaire with the maths module was carried out for the purpose of extending the sample basis for the TREE panel study on post-compulsory transitions (*Figure 1*). To this end, a slightly adapted version of the education monitoring module was implemented as a stand-alone web survey. [6] It was fielded immediately after the main ÜGK survey. The inclusion of the follow-up survey substantially broadened the sample base for all the questionnaire scales and instruments intended for the longitudinal monitoring of educational pathways.

This documentation describes all the scales incorporated in the ÜGK questionnaire and the estimation of the respective student scores. *The scales and scores documented here rely exclusively on the data collected in the main ÜGK survey and are intended for cross-sectional analyses of the ÜGK database.* By contrast, it is strongly recommended to use the scales based on the extended database including the follow-up survey (see *Figure 1*) for *longitudinal analyses* that use the ÜGK database along with the subsequent TREE panel waves. These scales for longitudinal purposes will be the subject of a separate documentation. [7]

[6] The interview mode and questionnaire are hence fully comparable with the respective ÜGK module. However, as the follow-up took place outside a proctored classroom environment, there may be setting-effects that reduce data comparability between the primary and the follow-up survey, although we do not assume this to be a major problem (see Gnambs & Kaspar, 2016).

[7] There is no reason to abstain from using the scales from the extended database for cross-sectional analysis as long as one is willing to assume that the data from the main field and from the follow-up [survey] are comparable. To check this assumption, the separate documentation will also include tests of measurement invariance across both survey-settings.

2. Selection and adaptation of scales for the ÜGK questionnaire

The ÜGK questionnaire incorporated a broad range of more than 90 item-based instruments from relevant research areas (listed in the appendix, starting on page 24). Instrument selection was motivated by criteria explained elsewhere. [8] It included primarily well-established scales that have been validated in previous research.

In addition, a *pretest survey* was implemented in spring 2015. The main goal was to test and improve mathematical competence assessments, the design of the context questionnaire and the fieldwork for the main survey. The pretest sample was split evenly across the three test languages Italian, French and German and included more than 2,000 students from 70 schools. One important pretest goal was to evaluate the measurement properties of the preliminary selection of questionnaire instruments and scales in the Swiss context. This included assessments of the dimensionality, reliability and the cross-language measurement invariance of the scales. Some of the scales had to be newly translated to make them available in all survey languages. In these cases, the pretest was used to check measurement invariance across language versions and to improve improper translations (see also footnote 9). Moreover, the pretest was used to clean up scales with dodgy items, to shorten others and, lastly, to narrow down and optimise the selection of instruments for the main survey. To ensure a comprehensive coverage of relevant concepts without unduly increasing the response burden and interview duration, it was necessary to shorten many scales to three or four items.

In order to preserve the measurement properties of the scales in our selection and to maximise data comparability, the original measurement instruments were implemented in the ÜGK questionnaire without modifications wherever possible. However, given the multitude of aspects to be considered in questionnaire construction (Dillman, Smyth & Christian, 2014), slight adaptations of the original instruments often could not be avoided.[9]

Due to the modular structure of the questionnaire, most ÜGK scales and associated scores are available either only for students who completed the mathematics version of the questionnaire or only for those who completed the 'education monitoring' version (i.e., for about half of the total sample of approximately 22,000 students).

---

[8] The selection of measurement instruments and the rationale behind it are documented in Hascher et al. (2015) and Hupka et al. (2015).

[9] The manifold methodological, empirical and substantive reasons for such adaptations included the following: At the methodological level, there was the need to adapt instruments that were originally developed for a different survey mode (de Leeuw, Hox & Dillman, 2008: 311f.) and to standardise the format of each type of question in order to reduce the response burden and improve understandability (Dillman, Smyth & Christian, 2014: 210f.). Empirically, the pretest in some instances uncovered insufficient cross-language measurement invariance, which suggested the need to check and, in some cases, improve the translations of the instruments. Finally, there was the requirement to closely replicate some of the instruments from the first TREE survey (TREE1).

The modifications of the original instruments can pertain to both the question format and wording of stimuli as well as to the response scales and sometimes even to the items. In most cases, however, they are minor so that a substantial impact on the measurement properties and comparability of the resulting scales seems unlikely. It should also be noted that, for similar reasons, many popular scales are far less standardised in survey practice than generally perceived. And in case of several circulating scale versions, the original version of the scale is not necessarily the most appropriate.

In principle, the scales selected for ÜGK are one-dimensional, that is to say, they are designed to measure *one* theoretical construct or latent dimension each.[10] However, some of the scales are composed of several sub-dimensions, each representing several facets of one overarching construct. As researchers may wish to empirically distinguish between the sub-dimensions of these scales, the scientific use file of ÜGK includes additional student scores for each sub-dimension. The following table shows both the main and sub-dimensions of the scales in question.

Table 1 Scales with sub-dimensions

Scale – Main Dimension	Variable Name [1)]	Sub-dimensions	Variable Name [1)]
Maths module scales			
Achievement Expectations Parents	[expectp_fs]	Achievement Expectations: Mother	[expectm_fs]
		Achievement Expectations: Father	[expectf_fs]
Instructivist Learning	[instreplearn_fs]	Instructivist Learning: Teachers' Instructions	[instrlearn_fs]
		Instructivist Learning: Repetitive Practice	[replearn_fs]
Social Learning	[soccomlearn_fs]	Social Learning: Social Arrangement	[soclearn_fs]
		Social Learning: Communication	[comlearn_fs]
System Aspect	[sysformasp_fs]	System Aspect: Logical Thinking	[systasp_fs]
		System Aspect: Formalism	[formasp_fs]
Teacher: Cognitive Activation [2)]	[cogself_fs]	Teacher: Cognitive Activation: Potential Solutions and Argumentation	[cogself1_fs]
		Teacher: Cognitive Activation: Strategies and Learning from Mistakes	[cogself2_fs]
Education monitoring module scales			
Global Self-Esteem [3)]	[sel_fs]	Positive Global Self-Esteem [4)]	[sele_fs]
		Negative Global Self-Esteem [4)]	[seld_fs]
Embodied Cultural Capital	[inccap_fs]	Embodied Cultural Capital: Manners	[manners_fs]
		Embodied Cultural Capital: Verbal Skills	[verbskill_fs]
Cultural Activities	[cult_fs]	Popular Cultural Activities	[cultlow_fs]
		Highbrow Cultural Activities	[culthigh_fs]

1) The names of the variables for the individual student scores in the ÜGK scientific use file are given in brackets. 2) As this scale is not one-dimensional in ÜGK, we distinguish two (inductively optimised) sub-dimensions. 3) In accordance with Huang et al. (2012) and Donnellan et al. (2016), this scale is clearly two-dimensional in ÜGK. 4) The sub-dimension labels were adopted from Huang et al. (2012).

The maths modules scales are displayed in the upper half of the table and the monitoring module scales in the lower half. As they rely on the same item sets, data analysts should take

[10] One should note, however, that the one-dimensionality of the selected scales may be empirically controversial. For one specific ÜGK scale, Global Self-Esteem (according to Rosenberg, 1979; 2014), we are aware that this is the case (see von Collani & Herzberg, 2003; Huang & Dong, 2012; Donnellan, Ackerman & Brecheen, 2016). With respect to this scale, we decided to provide the student scores for both the one-dimensional model and for the two sub-dimensions described in the literature. Hence, we treat this scale the same way as other scales with sub-dimensions and leave it up to the data users to decide on the appropriate scaling solution.

care to use *either* the student score for the main dimension *or* the scores for the subdimensions of a given scale.

Some of the instruments described in this documentation are based on two items only, making it impossible to fit any scaling model to the data. Henceforward, we call scores derived from these very short, item-based instruments below 'item-based composites' (see *Table 2* below).[11]

Table 2 Item-based composites [1)]

Concept Dimension	Variable Name [2)]	Number of Items
Big Five Inventory		
Extraversion	[big5_e_comp]	2
Agreeableness	[big5_a_comp]	3 [3)]
Conscientiousness	[big5_c_comp]	2
Neuroticism	[big5_n_comp]	2
Openness	[big5_o_comp]	2
Locus of Control		
Internal Locus of Control	[loc_i_comp]	2
External Locus of Control	[loc_e_comp]	2
Effort and Perseverance	[effper_comp]	2
Family Value Orientation	[vafa_comp]	2
Parents: Reading Interest	[joyreadp_comp]	2
Emotional Closeness to Parents	[closep_comp]	2

1) With the exception of 'Effort and Perseverance' (core questionnaire, full sample), all composites belong to the education monitoring module. 2) The names of the variables containing the composite scores in the ÜGK scientific use file are given in brackets. 3) For composites with one extra item, see Rammstedt and John (2007: 210).

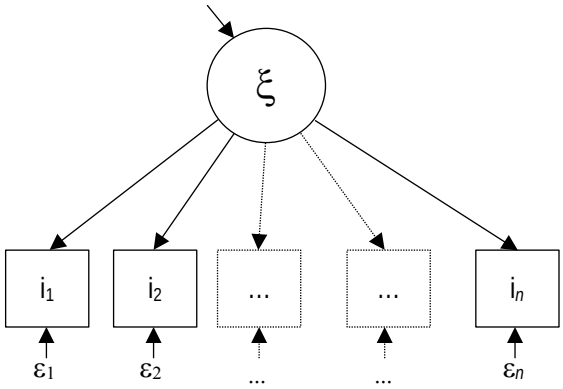
3. Statistical modelling

As mentioned above, the scales in the ÜGK questionnaire are item-based instruments intended to measure *one* theoretical construct each. Confirmatory factor analysis (CFA) is a common approach to the empirical estimation of latent (i.e., not directly observable) characteristics captured by such measurement instruments (see, e.g., Long, 1983; Schmitt, 2011). As our selection of scales is restricted to validated instruments that were designed to measure a common latent dimension, we limited ourselves to fitting a straightforward one-dimensional CFA model (see *Figure 2*, and Aichholzer, 2017: 80-84) to each scale-specific item set. The CFA model illustrated in *Figure 2* relies on n items (i_1, i_2, …, i_n) with associated item-level measurement errors ε_n, which all measure the same latent dimension ξ. For scales with several subdimensions (see *Table 1* above), a separate CFA model is fitted to each subdimension.[12]

[11] Student composite scores were calculated as the mean value of the (imputed) item ratings.

[12] An alternative approach would be to fit second-order CFA models to each dimension (Aichholzer, 2017: 89f.).

Figure 2 One-dimensional confirmatory factor model



For every model estimated hereafter, selected model parameters, fit statistics and scale quality measures are reported in the technical appendix (starting on page 24). This includes a test of one-dimensionality, various measures of internal scale consistency as well as tests and indices of cross-language measurement invariance. Throughout this documentation, our primary focus is the quality of the scales (and the corresponding student scores) rather than model fit. If the fit of the straightforward one-factor model turns out to be poor, we neither modify the model to improve fit nor do we test alternative (e.g., multi-dimensional) models. It is up to the data user to judge whether the one-dimensional CFA models are appropriate and whether the scales have the required properties.

3.1 *Estimation of the confirmatory factor models*

In its standard form, structural equation modelling – including CFA as a special case – relies on a number of quite restrictive assumptions, which are hardly ever met in practice. Basically, the observations should be independent, and the indicators should be measured on a continuous scale (interval-level measurement) and follow a multi-normal distribution (see, e.g., Hoyle, 2000). As regards the ÜGK database, none of these assumptions holds: The two-stage sampling procedure implies that observations are clustered within schools (see Verner & Helbling, forthcoming) and are hence not independent. Moreover, measurement of the indicators is at ordinal (or binary) level as it mostly relies on Likert-type rating scales. And last but not least, the skewed univariate distributions of many ratings are hardly consistent with the required multivariate normality.

The methodological literature offers a wide range of suggestions on how to relax some of the assumptions of the standard SEM model and how to deal with ordinal, binary or skewed indicators and clustered observations (cf., e.g., Bryant & Jöreskog, 2016).[13] In particular, the suggestions include two-stage estimation methods that exploit polychoric correlations and generalised structural equation models (GSEM) that are suited for categorical indicators (Rhemtulla, Brosseau-Liard & Savalei, 2012; Bryant & Jöreskog, 2016). However, there is

[13] Clustered observations may not only affect variance estimation and model fit but also bias the estimation of model parameters (i.e., factor loadings, cf. Stochl et al., 2016, Muthén & Satorra, 1995; Wu & Kwok, 2012).

currently no well-established, generally accepted estimation approach tailored to both ordinal indicators that are not normally distributed and a complex sample with clustered observations.

We therefore follow the recommendations of Rhemtulla et al. (2012; similarly Harpe, 2015: 843) regarding the accurate estimation of CFA models on the basis of ordinal, Likert-type indicators. They suggest two different estimation strategies depending on the length of the rating scales. For item responses that rely on a rating scale with at least five points (i.e., ordered discrete response categories), they suggest a two-step estimation based on polychoric correlations. For item evaluations that rely on shorter rating scales with four or less points, a generalised structural equation model (GSEM) is in order. Below, we describe these estimation strategies in more detail.[14] As our primary goal is to estimate accurate student scores, we also implement some sensitivity checks to assess the equivalence of student scores obtained via alternative model estimation strategies (see *section 3.2.1*).

3.1.1 *Two-step estimation based on polychoric inter-item correlations*

The two-step approach starts with the estimation of a matrix of polychoric correlations between all items of a given scale (tetrachoric correlations, respectively, in the case of dichotomous items).[15] In the second step, maximum likelihood estimation is used to fit the one-dimensional CFA model from *Figure 2* to the resulting correlation matrix.[16] The models are identified by setting the loading of the first item and the variance of the latent factor to one. The CFA models are also estimated separately for each of the three language subsamples. This allows for multi-group analysis designed to test and assess measurement invariance across the survey languages (see *section 4*, and, e.g., Steinmetz et al., 2008; Milfont & Fischer, 2015).

Below we briefly describe how we dealt with (a) the complex ÜGK sample and (b) with missing item values in the context of the two-step estimation approach.

(a) *Complex sample design and survey weighting*

The ÜGK survey relied on a random sample of students that was *disproportionally* stratified by cantons and type of cantonal curriculum (Verner & Helbling, forthcoming).[17] An unbiased estimation of any population characteristic requires the *application of an appropriate survey weight* to account for the disproportional design of the sample. This also pertains to the estimation of polychoric correlations or the parameters of the CFA models to be estimated (e.g., factor loadings).[18] We therefore applied the appropriate survey weights when estimating the polychoric correlations.

---

[14] All calculations were performed using Stata version 15.0.

[15] A polychoric correlation is defined as the maximum likelihood estimate of the correlation between two hypothetical, normally distributed continuous latent variables derived from two corresponding ordinal indicators. Estimation was done using the Stata-package 'polychoric' by Stas Kolenikov (from http://staskolenikov.net/stata).

[16] Maximum likelihood estimation has been found to be among the most appropriate estimation methods (together with ULS and DWLS, see Yang-Wallentin, Jöreskog & Luo, 2010) for analysing polychoric correlations derived from ordinal indicators.

[17] Lower secondary schools in Switzerland are mostly "tracked", i.e., students are enrolled in separate programmes with varying academic requirements.

[18] Weighting would only be unnecessary in case of a strict invariance of the postulated scaling model across population members of any kind. If this strong assumption were met, the damage of

It should be noted that the design of the two complementary random subsamples assigned to the distinct questionnaire modules (according to *Figure 1*) differs with regard to the shape of disproportional cantonal stratification.[19] Hence, whenever data from one questionnaire module is analysed, the subsample design should be taken into account by applying the appropriate survey weight for this purpose. When we estimate the polychoric correlations, we therefore have to select one out of three tailored survey weights, depending on whether the scale is embedded in the maths module, in the education monitoring module or in the core questionnaire.[20]

As regards the two-step estimation approach, it should be noted that variance estimation does not account for the clustering of observations within schools implied in the two-stage sampling (see Verner & Helbling, forthcoming).

(b) Handling of missing item values

Missing item values are not a major problem that affects the scales in the ÜGK survey. As in all scales, however, there is a small share of missing item values, owing mainly to item non-response. With the exceptions mentioned below, the share of cases with missing information on at least one item of the scale does not exceed 5%. For two out of three scales, the percentage is below 1%.

A considerably higher share of missing values results for half of the items of each of the four scales that measure different facets of 'specific self-efficacy' in mathematics. This is a direct consequence of the questionnaire design (and therefore not a matter of methodological concern[21]), as one half of the items of each of these scales was incorporated into the core questionnaire and the other half into the maths module. This implies that the share of missing item information is close to zero for the core questionnaire, whereas it rises to around 50% for the scales from one of the modules.

A relatively high share of missing values is also observed for two measures in which students evaluate the items on a rating scale that includes an explicit 'don't know' option. This pertains to the scale measuring 'perceived social network support' ('closupp_fs') and the two-item

unnecessarily applying survey weights would be limited to inflating the variances of the estimates to some degree (Bollen, Tueller & Oberski, 2013). Given the huge ÜGK sample, this would not be too disturbing.

[19] The reason is that the design of the two complementary subsamples has been optimised for two different purposes: the subsample drawn for the education monitoring module is geared toward maximising its analytical potential at the national level, whereas the maths module is tailored to providing the best results for separate analyses of cantons. In a nutshell, this was achieved by developing a disproportional subsampling scheme that further reinforces the general overrepresentation of small cantons among the subsample for the maths module and reduces it among the education monitoring subsample. The module-specific weights then correspond to the general survey weight for ÜGK (SUF variable 'smp_w_nrastubw') multiplied by the inverse of the within-canton subsampling fraction (cf. Verner & Helbling, forthcoming).

[20] The applied survey weights for the core questionnaire, the maths module and the education monitoring module are 'smp_w_nrastubw', 'smp_w_qmath' and 'smp_w_qtree', respectively (SUF variable names).

[21] Due to the randomised allocation of students to questionnaire modules, the missing-at-random assumption (MAR), which is crucial for the imputation of missing values, is almost perfectly met here.

composite for parents reading interests ('joyreadp_comp'). For both instruments, the share of missing information rises to about 8% when explicit don't-know answers are included.[22]

Finally, there are four instruments containing some items that could not be administered to a minor portion of the sample.[23] With one exception, the overall share of cases with at least one missing item does not exceed 5% in these cases.[24]

These special cases and exceptions notwithstanding, the fraction of missing items is low to very low for the bulk of the scales. Hence, the impact of missing item information is presumably limited.

We applied *multiple imputation* to cope with missing values when estimating the scaling models (Rubin, 1996; White, Royston & Wood, 2011). Basically, missing item information was imputed – scale-by-scale – on the basis of all valid items pertaining to the same scale. The imputed samples thus cover all cases with a valid response for at least one of the items of a given scale. Given the ordinal measurement level of the item ratings, we applied chained equations with an ordinal (or, in a few cases, binary) logit link to create samples with imputed values (Royston, 2011). Following the rules of thumb given in White et al. (2011: 388), we set the number of imputations to five.[25] For each imputed dataset, we separately calculated a matrix of polychoric correlations and combined it to estimate the CFA models.[26]

For each scale-specific CFA model, we calculated statistics and indices describing factor structures, model-fit and scale properties (see *section 5* and the technical appendix).

3.1.2 *Generalised structural equation model for categorical indicators (one-step estimation*)

The scales relying on short response scales with four or less points (including binary items) were analysed using a generalised structural equation model (GSEM), as recommended in the literature (Rhemtulla, Brosseau-Liard & Savalei, 2012; Bryant & Jöreskog, 2016). Model parameter estimates were derived in one step directly from the microdata through numeric

---

[22] Missing item values owing to explicit don't-know answers and item non-response were imputed together.

[23] Some items referring to specific relatives (e.g., the father) have not been administered when the students previously indicated that these relatives do not exist (this pertains to the instruments: Family Education Support, Parents Achievement Expectations, Parents Reading Interest and Emotional Closeness to Parents). The resulting missing values were treated the same way as other types of missing information. Although this is perhaps not an ideal solution in these cases, a substantial bias seems unlikely given the mostly very low number of cases that this applies to.

[24] The exception is the 'Family Educational Support' scale ('famdedsup_fs') for which the share of cases with at least one missing item amounts to nearly 16%. This owes mainly to the item tapping sibling support, which was not administered among students who previously indicated to have no siblings (see also footnote 23).

[25] The relatively low number of imputations seems appropriate for two additional reasons: First, we are primarily interested in unbiased point estimates of population parameters (e.g., factor loadings) and to a lesser degree in between-imputation and sampling variances. Second, some exploratory reproducibility checks as suggested by White et al. (2011: 387) indicate that the polychoric correlations and other point estimates are highly stable for an even smaller number of imputations.

[26] We simply average the matrices and do not consider between-imputation variance (see also footnote 25).

integration.[27] Contrary to the two-step approach, this amounts to a full-information, true maximum likelihood method (Bryant & Jöreskog, 2016: 192). We henceforth adopted the GSEM version of a one-dimensional CFA model, mostly with an ordinal logit link to account for the ordinal measurement level of the item sets to be analysed.[28]

(a) Accounting for the complex survey design

GSEM, as implemented in Stata, is able to account for complex sample designs. In particular, we used survey weights (as described in 3.1.1 a) to obtain unbiased population estimates of the model parameters and applied cluster-robust variance estimation, which controls for the clustering of students within schools. Still, we assume that there is no substantive variation in the measurement model across schools (cf. Wu & Kwok, 2012).

(b) Handling of missing item values

GSEM estimation proceeds on an equation-by-equation basis. In the context of a simple one-dimensional CFA model, this amounts to an implicit treatment (i.e., imputation) of missing item values, as each item is represented by a separate equation.

One drawback of the GSEM approach is that the calculation of most established statistics to describe model fit and scale properties is not straightforward. This is why we complemented the GSEM estimations for the item sets with short response scales by a separately estimated two-step model, as described in section 3.1. If the resulting factor structures and student scores do not substantially differ from those obtained via the GSEM approach, this may be taken as indirect evidence that the two-step approach works sufficiently well and its assumptions are met (in the appendix, we therefore also check for the equivalence of both types of student scores). Hence, the model and scale statistics taken from the two-step CFA model are likely to be valid approximations as well.

3.2 *ÜGK student scores*

3.2.1 *Calculation and robustness of student scores*

For instruments relying on items with rating scales of 5 or more points, the student scores in the scientific use file (and the related descriptive statistics in the appendix) represent *regression factor scores* (see StataCorp, 2017: 582f. for details) from the two-step CFA models described in *section 3.1.1*. For scales based on item sets with short response scales (four or less categories), the student scores in the SUF are *empirical Bayes means* based on the GSEM models (ibid.: 566). The *variable names assigned to the student scores* in the scientific use file ÜGK are mostly composed of the root of the variable names of the involved items and the suffix "_fs", which is used as a marker for all student score variables. The corresponding suffix for the item composites from *Table 2* is "_comp". The *variable labels assigned to the student scores* and item composites correspond to those contained in the scale-specific documentation in the appendix. For an unequivocal interpretation of the student scores in the SUF, it is recommended to inspect the factor loadings (see *section 4*). As a general rule,

[27] The integration method is a non-adaptive Gauss–Hermite quadrature with seven integration points (StataCorp, 2017: 562).

[28] The ordinal logit link reduces to a simple logit link for the two scales that include binary items.

however, a high factor score will indicate that students score high on the latent dimension that is designated by the variable label of the student score variable.

For all scales, the model, scale and test statistics reported in the appendix rely on the two-step estimation approach described in section 3.1.1. This explicitly also applies to those instruments based on short response scales, where the student scores (and the related factor-score descriptives in the appendix) are derived from a GSEM model. We also check the calculation of student scores for robustness by reporting the shared variance of both types of student scores (from SEM and GSEM) as measured by the coefficient of determination (CD). If their shared variance is close to 100% (i.e., CD approaches 1), one may safely conclude, first, that the different modelling strategies have a negligible impact on student scores and, second, that it also seems reasonable to take the various fit and scale statistics obtained from two-step estimation as good approximations. As documented scale by scale in the appendix, the coefficient of determination is indeed very close to 1 for most scales (> .94 for 41 out of 46 involved scales). There are five exceptions, however, in which the shared variance is substantially lower (between 72 and 90%), indicating that some of the additional assumptions needed for the two-step model have probably been violated. This pertains to the scales measuring absenteeism ('truancy_fs'), family wealth as mirrored in home possessions ('wealth_fs'), cultural activities including one of its subscales ('cult_fs','culthigh_fs') and students' maths self-concept ('matcon_fs'). For these scales, the model and scale statistics reported in the appendix should be interpreted with great caution, if at all. Still, this does not indicate that the student scores estimated via the GSEM approach are biased in any way.

For an additional robustness check for the student scores, we re-estimated the confirmatory factor models in one step directly from the student microdata using the MLMV method (StataCorp, 2017: 574). This allows us to control for the complex survey design through weighting and cluster-robust estimation and, at the same time, to implement an alternative full-information maximum-likelihood approach to account for missing item values. Let us again look at the shared variances between the student scores obtained via the MLMV method and those via the two-step approach described in section 3.1.1.[29] With the only exception of the aforementioned wealth scale, the shared variances uniformly exceed 96% (i.e., CD > .96) for all of the 82 scales in this documentation. This again may be taken as indirect evidence that the additional assumptions of the two-step approach regarding multivariate normal distributions and the measurement level are met and, hence, that the statistics and indices derived from it are valid. To sum up, the robustness checks imply that with the few exceptions mentioned above, student-score estimates are very robust across the three different estimation methods recommended for the type of data analysed here.[30]

[29] A disadvantage of this method is that many statistics needed to judge model-fit and scale qualities are unavailable.

[30] This may be due to the fact that we analyse short, one-dimensional scales based on a large sample.

3.2.2 *Brief remarks on the inclusion of student scores in multivariate statistical models*

Instead of using the scale-specific student scores, there are often good reasons to embed scale-specific CFA models into a more comprehensive structural equation model of substantive interest and to fit them together in one step (cf., e.g., Aichholzer, 2017). It should be noted, however, that simultaneous estimation of both the measurement and the substantive part of a structural equation model is not necessarily the best choice (cf. Devlieger & Rosseel, 2017): When one analyses a subsample of limited size, for instance, robust estimation of more complex models may be impossible. Moreover, even when the sample is large, misspecification bias in one part of a complex model may spread to other parts when they are fitted in one step. A two-step approach employing previously estimated factor scores to investigate the substantive part of the model may have methodological merits in this respect (ibid.). This approach also has methodological drawbacks, however, basically because it implicitly treats factor scores as error-free measures of the latent dimensions to be analysed.[31] Some of the resulting problems, possible biases and correction methods are discussed, for example, by Croon (2002), Lu and Thomas (2008), Jin et al. (2016), or Devlieger and Rossel (2017).

4 Content and interpretation of the reported scale-specific results

In this section, we outline the various statistics, indices and quality measures reported in the scale specific appendix. For each scale (or subscale, cf. *Table 2*), this report includes two pages with a variety of scale-specific statistics. Below, we take the scale that measures *'Parental Pressure to Achieve'* as an example to illustrate the scope and interpretation of scale-specific results. *Figure 3* displays the results for this scale as they appear in the appendix. If nothing else is mentioned, all reported results refer to the two-step estimation of the CFA model according to *Figure 2*. However, the student-scores descriptives refer to the scores obtained from the GSEM model, as the 'press' items are rated on a four-point scale (see section 3.2.1). The header of each scale-specific results section includes the name of the scale that is also used to label the related student-score variable in the scientific use file. In addition, the questionnaire module (cf. *Figure 1*) that incorporates the referenced scale is also documented in the header. It should be remembered that the sample size with valid information is approximately halved when a scale is embedded in one of the questionnaire modules (instead of in the core questionnaire).

On the upper left section of the first page of scale-specific results, we report a number of common model and fit statistics for the CFA model according to *Figure 2* (cf. Schreiber et al., 2006).[32] As pointed out in the introduction to section 3, we should bear in mind that model fit is neither a primary concern here nor easy to judge given the data at hand: First, many ÜGK scales rely on very short instruments with three items only, so that the fitted and saturated model are the same, making it impossible to assess model fit. Contrary to *Figure* 3, the

[31] A random extraction of plausible values from the posterior distributions of the CFA models could be a quite obvious solution to this. However, contrary to competence assessment, this is an uncommon approach in the scaling of questionnaire items, possibly because of the reduced convenience this entails for data analysis.

[32] Formal definitions are given in the technical part of the Stata manual (StataCorp, 2017: 577).

Figure 3 Example of the reported scale-specific results (first results page)

Scale: Parental Pressure to Achieve Monitoring Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	271	2	.000
	Baseline vs. Saturated	14102	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.112
	90% Confidence Interval: Lower Bound	.101
	90% Confidence Interval: Upper Bound	.124
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	97596
	Bayesian Information Criterion (BIC)	97683
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.981
	Tucker-Lewis Index (TLI)	.943
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.024
	Coefficient of Determination (CD)	.822

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.814
(Cronbach's alpha = .756)	
McDonald's Omega	.815

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.96
Factor 2	-.04
Factor 3	-.09
Factor 4	-.17

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
press1	0.68	.007	0.67	0.69
press2	0.70	.007	0.68	0.71
press3	0.79	.006	0.78	0.81
press4	0.72	.006	0.71	0.74

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
press1	2.2	1.0	1	4	10638
press2	3.0	0.9	1	4	10634
press3	3.1	0.8	1	4	10631
press4	2.8	0.9	1	4	10641

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
press1	1.61	-1.41	0.66	2.95
press2	1.79	-3.60	-1.78	0.78
press3	2.44	-5.21	-2.32	1.40
press4	1.94	-3.50	-1.23	1.55

model-fit statistics designed to capture discrepancies between model and observed data are therefore not informative when it comes to three-item scales. By definition, the likelihood-ratio test of the current against the saturated model, the RMSEA statistics, the baseline comparisons (CFI, TLI) and the size of the residuals (SRMR) will indicate a perfect fit (i.e., be either zero or one) for every three-item scale. Second, the very large size of the sample, even for the scales embedded in either one of the modules, is likely to result in even negligible deviations from the postulated model becoming highly significant that are of no substantive interest. Third, as the reported fit statistics do not take into account the clustering of observations within schools, all chi-square-based statistics are likely to be further inflated and

inferences on model fit in general may be biased (see, e. g., Stochl et al., 2016: 213).[33] Hence, all the fit statistics should be interpreted with caution, even for scales with at least four items.

The *model and fit statistics* reported include two likelihood-ratio tests as well as various common goodness-of-fit statistics, as discussed in the SEM literature (cf. Schreiber et al., 2006). The *likelihood-ratio tests* compare the current against the saturated model and the baseline model (basically postulating uncorrelated items), respectively. Ideally, we would expect a non-significant likelihood-ratio test of the current against the saturated model, which, for the reasons given above, is an unlikely result, however (see also van der Eijk & Rose, 2015). Moreover, for a well-fitting model, we expect the *comparative fit index (CFI)* and the *Tucker-Lewis index (TLI)* to approach 1, whereas the *root mean square error of approximation (RMSEA)* and the *standardised root mean squared residual (SRMR)* should be close to 0. Conventional cut-off criteria indicating a good fit between the hypothesised model and the observed data are ≥ .95 for CFI and TLI, ≤ .06 for RMSEA and ≤ .08 for SRMR (see Hu & Bentler, 1999). Regarding *Figure 3*, one could tentatively conclude that the one-dimensional CFA model fits the achievement-pressure scale sufficiently well, with the reservation that the RMSEA clearly exceeds the mentioned cut-off. Two fit measures designed to compare different models, the *Akaike's and Bayesian information criteria (AIC, BIC)*, are also reported. They may serve as a point of reference if data users wish to fit alternative scaling models to the ÜGK data. Finally, the *coefficient of determination (CD)* may be considered as an alternative measure of composite reliability (in the sense of internal consistency, cf. Bollen, 1989: 220f.), to be interpreted similarly to the reliability measures below.

The output section to the right of the model-fit statistics presents the results on *scale reliability and dimensionality*. Among the various conceptualisations of measurement reliability discussed in the literature (e.g., Bollen, 1989), *internal scale consistency* is the most widely used in practical research. One important reason for this is certainly that internal consistency may be easily assessed without additional re-test or parallel measurements of the indicators. It should also be noted, however, that consistency measures avoid several conceptual drawbacks of possible alternatives (see Bollen, 1989: 209ff.). We report three alternative measures of internal scale consistency: *Cronbach's alpha* is still the most widespread, although much criticised, consistency measure (ibid.: 217, Sijtsma, 2009; Revelle & Zinbarg, 2009; Trizano-Hermosilla & Alvarado, 2016). In a nutshell, it is widely recognised that alpha underestimates internal consistency if the indicators are ordinal or congeneric (i.e., not tau-equivalent) as is typical of most practical research situations. We nevertheless do report the classical version of alpha as it is part of most survey documentations and – if interpreted as a lower-bound estimate of internal scale consistency – may still be useful for comparative purposes.[34] In addition, we also report *ordinal Cronbach's alpha*, which is calculated the same way as classical alpha but from the matrix of polychoric instead of Pearson correlations (see Gadermann, Guhn & Zumbo, 2012: 5). This avoids downward bias due to the ordinal measurement. Finally, we also report *McDonald's omega*, which is one of the most recommended measures of internal consistency. Omega is calculated based on the factor loadings of the one-dimensional CFA model (according to formula 1 in Trizano-Hermosilla &

[33] Results by Stochl et al. (2016: 213, Figure 6) suggest that the root mean square error of approximation (RMSEA) is the perhaps most robust index of model fit in this situation. Conventionally, a value below .05 is taken as an indication of a good fit.

[34] The Stata package Alphawgt, which allows for weights, was used to calculate alpha (Jann, 2004).

Alvarado, 2016), which implies that it is adjusted for ordinal measurement. As omega is appropriate for congeneric indicators, it is probably the most adequate measure overall of internal scale consistency in our context (see also Yang & Green, 2015). Basically, values close to 1 indicate high internal consistency for all three measures. Looking at *Figure* 3, many researchers would probably interpret the almost identical ordinal alpha and omega values of .814 and .815, respectively, as an indication of a 'good', consistent scale. It should be noted, however, that the widely used rules of thumb to determine whether internal scale consistency can be considered 'acceptable' or 'good' (usually values above .7 and .8, respectively) are not without problems. First, there exist various variants of such rules of thumb with different critical thresholds. Second, and more importantly, such rules should not be blindly applied, as the acceptable level of internal consistency depends strongly on the type of analysis to be performed (Lance, Butts & Michels, 2006).[35]

A crucial assumption of the estimated CFA models is that the analysed item set captures only one latent construct. Therefore, we have also included a *test of the assumed one-dimensionality.* However, assessing dimensionality of Likert-type items is quite 'risky business', as van der Eijk and Rose put it (2015). We used explorative factor analysis of polychoric correlations followed by Horn's parallel analysis to assess the dimensionality of the item sets, which proves to be a comparatively well-performing method (ibid.; Garrido, Abad & Ponsoda, 2013).[36] Basically, we applied an eigenvalue criterion that was corrected for random factors to account for sampling variance to determine the number of factors to be retained. In *Figure 3*, this approach gives us no reason to believe that the achievement-pressure scale is not one-dimensional, as only the eigenvalue of the first factor exceeds the critical value of zero. If we leave aside that the scales are composed of several sub-dimensions (cf. *Table 1*), the eigenvalues of the second factor are mostly below or only very slightly above 0 for most of the scales in this documentation. This being the case, we have no clear indication that the one-dimensionality assumption is violated.

The section below the model-fit statistics in *Figure 3* documents the *standardised factor loadings* for each item, including standard errors and the confidence intervals. The item names correspond to those in the scientific use file. High standardised loadings above – say, .7 or .8 – indicate that neither measurement errors nor strong unique factors contribute substantially to the variance of the observed indicators. To the right of the loadings, a number of *item descriptives* are reported, including the mean, the standard deviation, the range of the rating scale applied for item evaluation (min., max.) and the number of students with valid item ratings (see section 3.1.1 b for details). This section also includes measures of factor-score equivalence across different estimation methods (see section 3.2.1 for details).

At the bottom of the first page of scale-specific results, we report the *parameters of the categorical GSEM model* (cf. section 3.1.2), where estimated. Note that for this model, there are two-types of item-specific parameters, namely, path coefficients ('coef') that measure the effect of the latent variable on the indicator rating, and the estimated cut points ('cutx') on the logit distribution that separate the rating scale category 1 from category 2, category 2 from

[35] There are various, somewhat dubious rules of thumb that distinguish different levels of internal scale consistency (i.e., Cronbach's alpha). A popular variant is (see en.wikipedia.org/wiki/Cronbach's-_alpha, accessed on 16 April 2018): $\alpha < .5$: unacceptable; $.5 \leq \alpha < .6$: poor; $.6 \leq \alpha < .7$: questionable; $.7 \leq \alpha < .8$: acceptable; $.8 \leq \alpha < .9$: good; $.9 \leq \alpha$: excellent.

[36] The parallel analysis relies on the user-written 'paran'-package (Dinno, 2009).

category 3 and so on. Hence, the number of estimated 'cut' parameters equals the number of ordered rating categories minus one. Remember that the GSEM model was used to generate student scores (see section 3.1) in those cases in which students' item evaluations relied on short rating scales with four or less points (as documented by the item descriptives).

A second page of scale-specific results (see *Figure 4* below) is dedicated to tests and indices assessing *measurement invariance across the three survey languages*. This is an important facet of measurement quality, as student scores obviously should be comparable – i.e., measure the same concepts on a possibly invariant scale – across all kinds of subsamples of the underlying student population. Comparability across survey languages is an especially important and critical aspect of this. We focus on some of the most crucial tests suggested in the literature on the multi-group analysis of measurement invariance (e.g., Vandenberg & Lance, 2000; Milfont & Fischer, 2015) to assess cross-language measurement equivalence. On top of the results page, we first report a chi-square test of the *equality of the item-covariance matrices* across the three subsamples that received different language versions of the questionnaire.[37] If the hypothesis of equal covariance matrices is not rejected, this would be a strong indication of measurement invariance, making any further tests obsolete (ibid.).

The three chi-square tests assembled in the section below refer to the one-dimensional CFA model from *section 3.1.1*, which was re-estimated separately for each language subsample. Hence, the tests assume that a common latent dimension exists, and its invariance is investigated by means of multi-group analysis. The three tests are designed to distinguish different levels of measurement equivalence, as discussed in the literature (ibid.). The first test is for *metric measurement invariance*, that is, for equal factor loadings. A non-significant test indicates that there is no evidence against the postulated invariance of the factor loadings across survey languages. The second test takes the model with invariant loadings as its baseline and tests it against an alternative model with invariant loadings *and* intercepts, which implies *strong measurement invariance*. Third and lastly, the latter model is tested against an alternative positing *strict measurement invariance*, which furthermore requires invariant error variances (ε_i in Figure 2). Given the nested structure of the compared models, strong invariance would require that the first two tests are not significant, and strict invariance that all three tests are not significant. Although this is a rather standard approach to assess measurement equivalence, the reservations against chi-square-based fit statistics discussed above in conjunction with model fit also extend to chi-square-based multi-group comparisons: Even if the cross-language variations in the model parameters are negligible, these tests will almost always be significant given the huge sample analysed here. That is to say, a level of measurement equivalence that would be adequate for nearly all practical research purposes would still not be enough to pass these tests.

[37] Technically, this was done by specifying a multi-group model without a latent dimension and then testing a completely unconstrained against a constrained model with equal variances and inter-item covariances.

Figure 4 Example of the reported scale-specific results (second results page)

Scale: Parental Pressure to Achieve (continued) Monitoring Module

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1159	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	26	6	.000
Strong Invariance (Plus Equal Intercepts)	590	6	.000
Strict Invariance (Plus Equal Error Variances)	48	6	.000

Configural Factor Similarity	
Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.991
Italian vs. German Language Version	.986

Factor Score Equivalence: Group Specific vs. Invariant Models	
Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.959

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
press_fs	0.0	0.9	-2.4	1.7	10668

Share of Cases with Imputed Missing Values: 0.5%
(Equivalence of Scores from Robust MLMV: CD = .997)
(Equivalence of Scores from Two-Step Approach: CD = .984)

Therefore, we also report two additional measures of factor equivalence, which perhaps do better in meeting the practical needs of many data analysts. The first one, *Tucker's congruence coefficient (TCC)*, is a measure of *configural factor invariance* (calculated according to formula 1 in Lorenzo-Seva & ten Berge, 2006). Basically, it is a pattern-similarity measure that approaches 1 when the loading patterns observed in two groups or conditions are identical. We report the coefficient separately for each pair of survey languages. According to Lorenzo-Seva and ten Bergen (ibid.: 61), two factors may be considered as approximately equal for practical purposes if TCC exceeds .95. If we look at the scales documented in the appendix, this criterion is always met for all pairwise language comparisons. In addition, we also assess the degree of *micro-level factor equivalence at the level of student scores*. For this, we compare the student scores taken from an unconstrained model fitted separately for each language subsample with the student scores taken from a model for the entire sample on the assumption of strong measurement invariance (i.e., equal loadings and intercepts). If the differences between the former and latter are negligible across all language subsamples, this is a strong indication that – from a practical point of view – the measurement can be regarded

as sufficiently invariant. As a measure of micro-level agreement, we report – separately for each of the subsamples delineated by survey language – the *coefficient of determination (CD)*, which is calculated by regressing the student scores from the strong-invariance model on those from the unconstrained language-specific models. Where the CD indicates that both scores share, say, 98 per cent of their variance (i.e., CD ≥ .98), deviations from the postulated strong invariance model may be regarded as negligible. It should be noted that a perfect agreement cannot be expected even if the 'true' measurement model would be absolutely invariant as the estimated student scores also include some random error. This is particularly true for the scores gained through the separate analysis of the Italian questionnaire, where the already rather small subsample (*n* = 755) was approximately halved in most cases due to the modularisation of the questionnaire (cf. *Figure 1*).[38] Notably for the Italian questionnaire, the sampling errors in the factor loadings and hence also in the student scores are likely to be substantial.[39] With this in mind, one could perhaps also accept a coefficient of determination of, say, .95 as an indication of a still fair level of measurement equivalence. In case of the achievement-pressure scale in *Figure 3*, we might still conclude that the measurement is approximately equivalent in all three survey languages.

In the section below the measurement invariance tests and indices, we report the names ('press_fs' in *Figure* 3) and some descriptive statistics of the student-scores variables as stored in the scientific use file (from either ML-SEM or GSEM, depending on the length of the rating scales; see section 3.2.1). This also includes the total number of cases with available student scores and the share of cases with at least one missing item with an imputed value. Depending on the length of the rating scales applied for item evaluation, one or two measures of factor-score equivalence across different estimation methods are reported (see section 3.2.1).

[38] The samples for the French (*n* = 5'235) and the German (*n* = 16'349) versions are far larger (core questionnaire).

[39] In combination with skewed item distributions, this is probably also the reason why a few of the models underlying the invariance tests did not converge so that the subsamples for the French and the Italian languages had to be collapsed for this purpose. We added an explanatory note at the end of the measurement-equivalence section in the appendix, which is shaded in grey in these cases (e.g., the 'school reluctance' scale).

Some practical guidelines for using the ÜGK scales

It is up to the data users to check whether a scale shows the measurement properties required for their analysis. For each scale implemented in the ÜGK questionnaire, this documentation provides a selection of relevant statistics and measures (see appendix). *Section 4* of the introductory text includes brief explanations of the definitions and interpretations of the reported measures.

When using the ÜGK scales, the following remarks should be observed:

- The documented scales and student scores are *exclusively intended for cross-sectional analysis* of the ÜGK 2016 database. For *longitudinal analysis*, we strongly recommend using the scale versions constructed specifically for this purpose. This refers to the scales of the education monitoring module of the ÜGK questionnaire (see *Figure 1*): For these scales, student scores for a considerably enlarged sample will be available, as the sample base was enlarged by a complementary survey carried out by TREE (see section 1 for details). The respective scale versions will be described in a separate scale documentation.
- The reported scale-specific measures focus primarily on *reliability* (in the sense of internal consistency) and *measurement invariance* across survey languages. Contrary to reliability, *scale validity* is not addressed in this documentation. As ÜGK mostly implements commonly accepted, well-established scales, validity is not likely to be a major problem. In addition, the ÜGK database offers researchers many opportunities to conduct external validations tailored to their specific analytical needs.
- While all the ÜGK scales are intended to measure only one theoretical dimension, some of them include *several sub-dimensions* (see section 2, *Table 1*). For each of these scales, the scientific use file includes a student score for the main dimension as well as for each of the sub-dimensions. *As both types of scores were extracted from the same items, researchers should avoid employing both of them simultaneously within the same multivariate model*.
- Before making use of *student scores in the context of multivariate models*, the reader should take note of the brief remarks on this issue in section 3.2.2. Some scores represent item composites rather than scale values (cf. *Table 2*), which may, however, be used similarly. The names and labels of all items, student scores and composite variables in the technical appendix are the same as in the scientific use file of the ÜGK.
- In the process of estimating the confirmatory factor models and calculating the student scores, all *missing item information* was imputed – provided that at least one item of a given scale had a valid rating (see section 3.1.1 b for details).

References

Aichholzer, Julian (2017). *Einführung in lineare Strukturgleichungsmodelle mit Stata*. Springer: Wiesbaden (1. Auflage).

Bollen, Kenneth A. (1989). *Structural Equation with Latent Variables*. Wiley: New York.

Bollen, Kenneth, Stephen Tueller & Daniel Oberski (2013). 'Issues in the Structural Equation Modeling of Complex Survey Data.' International Statistical Institute (Ed.): *Proceedings of the 59th World Statistics Congress*: Hongkong.

Bryant, Fred B. & Karl G. Jöreskog (2016). 'Confirmatory Factor Analysis of Ordinal Data Using Full-Information Adaptive Quadrature'. *Australian & New Zealand Journal of Statistics* 58 (2): 173–196.

Croon, Marcel A. (2002). 'Using Predicted Latent Scores in General Latent Structure Models', pp. 195–224, in George A. Marcoulides & Irini Moustaki (Eds.): *Latent Variable and Latent Structure Models.* Lawrence Erlbaum.

de Leeuw, Edith, Joop Hox & Don. A. Dillman (2008). 'Mixed-Methods Surveys: When and Why', pp. 299–316, in Edith de Leeuw, Joop Hox & Don A. Dillman: *International Handbook of Survey Methodology.* Lawrence Erlbaum: New York.

Devlieger, Ines & Yves Rosseel (2017). 'Factor Score Path Analysis'. *Methodology* 13 (Supplement 1): 31–38.

Dillman, Don. A., Jolene D. Smyth & Leah Melani Christian (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. John Wiley & Sons: Hoboken (4th Edition).

Dinno, Alexis (2009). 'Implementing Horn's Parallel Analysis for Principal Component Analysis and Factor Analysis'. *Stata Journal* 9 (2): 291–298.

Donnellan, M. Brent, Robert A. Ackerman & Courtney Brecheen (2016). 'Extending Structural Analyses of the Rosenberg Self-Esteem Scale to Consider Criterion-Related Validity: Can Composite Self-Esteem Scores Be Good Enough?' *Journal of Personality Assessment* 98 (2): 169–177.

FORS (2018). Codebook ÜGK 2016. Manuscript in preparation.

Gadermann, Anne M., Martin Guhn & Bruno D. Zumbo (2012). 'Estimating Ordinal Reliability for Likert-Type and Ordinal Item Response Data: A Conceptual, Empirical, and Practical Guide'. *Practical Assessment, Research & Evaluation* 17 (3): 1–13.

Garrido, Luis Eduardo, Francisco José Abad & Vicente Ponsoda (2013). 'A New Look at Horn's Parallel Analysis with Ordinal Variables'. *Psychological Methods* 18 (4): 454–474.

Gnambs, Timo & Kai Kaspar (2016). 'Socially Desirable Responding in Web-Based Questionnaires: A Meta-Analytic Review of the Candor Hypothesis'. *Assessment* 24 (6): 746–762.

Harpe, Spencer E. (2015). 'How to Analyze Likert and Other Rating Scale Data'. *Currents in Pharmacy Teaching and Learning* 7 (6): 836–850.

Hascher, Tina, Christian Brühwiler, Andrea Erzinger, Boris Girnat et al. (2015). *Erläuterungen zu den Skalen des Kontextfragebogens Mathematikteil: Theoretischer Hintergrund und Forschungsinteressen.* Universität Bern, Pädagogische Hochschule St. Gallen & Pädagogische Hochschule FHNW (Hg.): Bern.

Hoyle, Rick H. (2000). 'Confirmatory Factor Analysis', pp. 465-497, in Steven D. Brown (Ed.): *Handbook of Applied Multivariate Statistics and Mathematical Modeling.* Academic Press: San Diego.

Hu, Litze & Peter M. Bentler (1999). 'Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives'. *Structural Equation Modeling* 6 (1): 1–55.

Huang, Chiungjung J. & Nianbo B. Dong (2012). 'Factor Structures of the Rosenberg Self-Esteem Scale: A Meta-Analysis of Pattern Matrices'. *European Journal of Psychological Assessment* 28 (2): 132–138.

Hupka-Brunner, Sandra, Ben Jann, Thomas Meyer, Christian Imdorf et al. (2015). *Erläuterungen zum Kontextfragebogen der ÜGK 2016: Allgemeiner Teil.* Universität Bern: Bern.

Jann, Ben (2004). *ALPHAWGT: Stata Module to Compute Cronbach's Alpha for Weighted Data*. Statistical Software Components S4444101). Boston: Boston College Department of Economics.

Jin, Shaobo, Hao Luo & Fan Yang-Wallentin (2016). 'A Simulation Study of Polychoric Instrumental Variable Estimation in Structural Equation Models'. *Structural Equation Modeling* 23 (5): 680–694.

Lance, Charles E., Marcus M. Butts & Lawrence C. Michels (2006). 'The Sources of Four Commonly Reported Cutoff Criteria: What Did They Really Say?' *Organizational Research Methods* 9 (2): 202–220.

Long, J. Scott (1983). *Confirmatory Factor Analysis*. Vol. 33 of 'Sage University Paper Series on Quantitative Applications in the Social Science', edited by Michael S. Lewis-Beck. Sage: Newbury Park.

Lorenzo-Seva, Urbano & Jos M. F. ten Berge (2006). 'Tucker's Congruence Coefficient as a Meaningful Index of Factor Similarity'. *Methodology* 2 (2): 57–64.

Lu, Irene R. R. & D. Roland Thomas (2008). 'Avoiding and Correcting Bias in Score-Based Latent Variable Regression with Discrete Manifest Items'. *Structural Equation Modeling* 15 (3): 462–490.

Milfont, Taciano L & Ronald Fischer (2015). 'Testing Measurement Invariance across Groups: Applications in Cross-Cultural Research'. *International Journal of Psychological Research* 3 (1): 111–130.

Muthén, Bengt O. & Albert Satorra (1995). 'Complex Sample Data in Structural Equation Modeling'. *Sociological Methodology* 25: 267–316.

Rammstedt, Beatrice & Oliver P. John (2007). 'Measuring Personality in One Minute or Less: A 10-Item Short Version of the Big Five Inventory in English and German'. *Journal of Research in Personality* 41 (1): 203–212.

Revelle, William & Richard E. Zinbarg (2009). 'Coefficients Alpha, Beta, Omega, and the GLB: Comments on Sijtsma'. *Psychometrika* 74 (1): 145–154.

Rhemtulla, Mijke , Patricia E. Brosseau-Liard & Victoria Savalei (2012). 'When Can Categorical Variables Be Treated as Continuous? A Comparison of Robust Continuous and Categorical SEM Estimation Methods under Suboptimal Conditions'. *Psychological Methods* 17 (3): 354–73.

Rosenberg, Morris (1979). *Conceiving the Self*. Basic Books: New York.

Rosenberg, Morris (2014). 'Self-Esteem Scale'. GESIS (Hg.): *Zusammenstellung sozialwissenschaftlicher Items und Skalen*: Mannheim. ZIS-Version 16.00. DOI: 10.6102/zis46.

Royston, Patrick (2011). 'Multiple Imputation by Chained Equations (MICE): Implementation in Stata'. *Journal of Statistical Software* 45 (4): 1–20.

Rubin, Donald B. (1996). 'Multiple Imputation After 18+ Years'. *Journal of the American Statistical Association* 91 (434): 473–489.

Schmitt, Thomas A. (2011). 'Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis'. *Journal of Psychoeducational Assessment* 29 (4): 304–321.

Schreiber, James B., Amaury Nora, Frances K. Stage, Elizabeth A. Barlow et al. (2006). 'Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review'. *The Journal of Educational Research* 99 (6): 323–338.

Sijtsma, Klaas (2009). 'On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha'. *Psychometrika* 74 (1): 107–120.

StataCorp (2017). *Stata Structural Eqation Modeling Reference Manual. Release 15.* College Station, Texas: Stata Press.

Steinmetz, Holger, Peter Schmidt, Andrea Tina-Booh, Siegrid Wieczorek et al. (2008). 'Testing Measurement Invariance Using Multigroup CFA: Differences between Educational Groups in Human Values Measurement'. *Quality & Quantity* 43 (4): 599–616.

Stochl, Jan, Peter B. Jones, Jesus Perez, Golam M. Khandaker et al. (2016). 'Effects of Ignoring Clustered Data Structure in Confirmatory Factor Analysis of Ordered Polytomous Items: A Simulation Study Based on PANSS'. *International Journal of Methods in Psychiatric Research* 25 (3): 205–219.

Trizano-Hermosilla, Italo & Jesús M. Alvarado (2016). 'Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements'. *Frontiers in Psychology* 7. Published online 2016 May 26: doi: 10.3389/fpsyg.2016.00769.

van der Eijk, Cees & Jonathan Rose (2015). 'Risky Business: Factor Analysis of Survey Data – Assessing the Probability of Incorrect Dimensionalisation'. *PLoS ONE* 10 (3). doi: 10.1371/journal.pone.0118900.

Vandenberg, Robert J. & Charles E. Lance (2000). 'A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research'. *Organizational Research Methods* 3 (1): 4–70.

Verner, Martin & Laura A. Helbling (in Vorbereitung). *Sampling ÜGK 2016: Technischer Bericht zu Stichprobendesign, Gewichtung und Varianzschätzung bei der Überprüfung der des Erreichens der Grundkompetenzen 2016.* Institut für Bildungsevaluation (IBE): Zürich.

von Collani, Gernot & Philipp Yorck Herzberg (2003). 'Zur internen Struktur des globalen Selbstwertgefühls nach Rosenberg'. *Zeitschrift für Differentielle und Diagnostische Psychologie* 24 (1): 9–22.

White, Ian R., Patrick Royston & Angela M. Wood (2011). 'Multiple Imputation Using Chained Equations: Issues and Guidance for Practice'. *Statistics in Medicine* 30 (4): 377–399.

Wu, Jiun-Yu & Oi-man Kwok (2012). 'Using SEM to Analyze Complex Survey Data: A Comparison between Design-Based Single-Level and Model-Based Multilevel Approaches'. *Structural Equation Modeling* 19 (1): 16–35.

Yang-Wallentin, Fan, Karl G. Jöreskog & Hao Luo (2010). 'Confirmatory Factor Analysis of Ordinal Variables with Misspecified Models'. *Structural Equation Modeling* 17 (3): 392–423.

Yang, Yanyun & Samuel B. Green (2015). 'Evaluation of Structural Equation Modeling Estimates of Reliability for Scales with Ordered Categorical Items'. *Methodology* 11 (1): 23–34.

SCALE APPENDIX

TABLE OF CONTENTS

Scales Related to the Acquisition of Mathematics Skills

Scales and Composites Related to School-to-Work Transitions

Scale: Reality-Based Learning

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	129	2	.000
	Baseline vs. Saturated	14527	6	.000
2)	**Root Mean Squared Error (RMSEA)**			.076
	90% Confidence Interval: Lower Bound			.065
	90% Confidence Interval: Upper Bound			.087
	Probability RMSEA <= 0.05			.000
3)	**Akaike's Information Criterion (AIC)**			145766
	Bayesian Information Criterion (BIC)			145853
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			.991
	Tucker-Lewis Index (TLI)			.974
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.016
	Coefficient of Determination (CD)			.832

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.807
(Cronbach's alpha = .779)	
McDonald's Omega	.811

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.94
Factor 2	-.04
Factor 3	-.11
Factor 4	-.15

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
realref1	0.61	.007	0.60	0.63
realref2	0.65	.007	0.64	0.66
realref3	0.80	.005	0.79	0.81
realref4	0.80	.005	0.79	0.81

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
realref1	3.8	1.5	1	6	11042
realref2	3.9	1.4	1	6	10995
realref3	3.7	1.5	1	6	10984
realref4	4.1	1.5	1	6	11035

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	388	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	210	6	.000
Strong Invariance (Plus Equal Intercepts)	116	6	.000
Strict Invariance (Plus Equal Error Variances)	78	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.983
French vs. Italian Language Version	.993
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.999
Language: French	.989
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
realref_fs	0.0	0.8	-2.1	1.6	11063

Share of Cases with Imputed Missing Values: 1.1%

(Equivalence of Scores from Robust MLMV: CD = .998)

Scale: Discovery / Exploratory Learning

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	132	2	.000
	Baseline vs. Saturated	19790	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.076
	90% Confidence Interval: Lower Bound	.066
	90% Confidence Interval: Upper Bound	.088
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	143687
	Bayesian Information Criterion (BIC)	143775
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.993
	Tucker-Lewis Index (TLI)	.980
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.013
	Coefficient of Determination (CD)	.867

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.858
(Cronbach's alpha = .836)	
McDonald's Omega	.859

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.30
Factor 2	-.06
Factor 3	-.09
Factor 4	-.13

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
disclearn1	0.73	.005	0.72	0.74
disclearn2	0.84	.004	0.83	0.85
disclearn3	0.81	.004	0.80	0.82
disclearn4	0.72	.005	0.71	0.74

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
disclearn1	3.5	1.6	1	6	11049
disclearn2	3.5	1.5	1	6	10986
disclearn3	3.6	1.5	1	6	11002
disclearn4	3.7	1.5	1	6	11006

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	712	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	83	6	.000
Strong Invariance (Plus Equal Intercepts)	126	6	.000
Strict Invariance (Plus Equal Error Variances)	190	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.985
French vs. Italian Language Version	.992
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.993
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
disclearn_fs	0.0	1.1	-2.3	2.1	11067

Share of Cases with Imputed Missing Values: 1.1%

(Equivalence of Scores from Robust MLMV: CD = .998)

Scale: Social Learning

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	5090	9	.000
	Baseline vs. Saturated	36459	15	.000

2)	**Root Mean Squared Error (RMSEA)**	.226
	90% Confidence Interval: Lower Bound	.221
	90% Confidence Interval: Upper Bound	.231
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	211536
	Bayesian Information Criterion (BIC)	211668
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.861
	Tucker-Lewis Index (TLI)	.768
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.096
	Coefficient of Determination (CD)	.912

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.869
(Cronbach's alpha = .849)	
McDonald's Omega	.865

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.20
Factor 2	.48
Factor 3	-.06
Factor 4	-.08
Factor 5	-.09
Factor 6	-.13

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
comlearn1	0.54	.007	0.52	0.55
comlearn2	0.51	.008	0.50	0.53
comlearn3	0.62	.006	0.61	0.64
soclearn1	0.83	.004	0.83	0.84
soclearn2	0.88	.003	0.87	0.89
soclearn3	0.87	.003	0.87	0.88

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
comlearn1	3.8	1.5	1	6	11035
comlearn2	3.5	1.5	1	6	11009
comlearn3	3.7	1.5	1	6	10993
soclearn1	4.0	1.6	1	6	11039
soclearn2	4.3	1.5	1	6	11004
soclearn3	4.2	1.5	1	6	10990

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	580	54	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	55	10	.000
Strong Invariance (Plus Equal Intercepts)	202	10	.000
Strict Invariance (Plus Equal Error Variances)	155	10	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.997
Italian vs. German Language Version	.997

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.998

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
soccomlearn_fs	0.0	0.8	-1.9	1.2	11065

Share of Cases with Imputed Missing Values: 1.2%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Social Learning: Social Arrangement

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	21585	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			100479
	Bayesian Information Criterion (BIC)			100545
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.914

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.904
(Cronbach's alpha = .882)	
McDonald's Omega	.905

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.16
Factor 2	-.07
Factor 3	-.11

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
soclearn1	0.85	.003	0.84	0.86
soclearn2	0.92	.003	0.92	0.93
soclearn3	0.84	.004	0.84	0.85

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
soclearn1	4.0	1.6	1	6	11039
soclearn2	4.3	1.5	1	6	11004
soclearn3	4.2	1.5	1	6	10990

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	142	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	25	4	.000
Strong Invariance (Plus Equal Intercepts)	54	4	.000
Strict Invariance (Plus Equal Error Variances)	21	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.999
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.999

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
soclearn_fs	0.0	1.2	-2.9	1.7	11060

Share of Cases with Imputed Missing Values: 1.0%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Social Learning: Communication

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	9617	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	111136
	Bayesian Information Criterion (BIC)	111202
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.816

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.782
(Cronbach's alpha = .751)	
McDonald's Omega	.786

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.47
Factor 2	-.10
Factor 3	-.18

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
comlearn1	0.70	.007	0.69	0.72
comlearn2	0.66	.007	0.65	0.68
comlearn3	0.85	.007	0.84	0.87

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
comlearn1	3.8	1.5	1	6	11035
comlearn2	3.5	1.5	1	6	11009
comlearn3	3.7	1.5	1	6	10993

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	261	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	9	4	.070
Strong Invariance (Plus Equal Intercepts)	53	4	.000
Strict Invariance (Plus Equal Error Variances)	17	4	.002

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.999
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
comlearn_fs	0.0	0.9	-2.1	1.8	11062

Share of Cases with Imputed Missing Values: 1.0%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Instructivist Learning

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	4517	20	.000
	Baseline vs. Saturated	29913	28	.000

2)	**Root Mean Squared Error (RMSEA)**	.143
	90% Confidence Interval: Lower Bound	.139
	90% Confidence Interval: Upper Bound	.146
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	286311
	Bayesian Information Criterion (BIC)	286487
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.850
	Tucker-Lewis Index (TLI)	.789
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.066
	Coefficient of Determination (CD)	.848

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.841
(Cronbach's alpha = .818)	
McDonald's Omega	.842

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.18
Factor 2	.36
Factor 3	.21
Factor 4	.05
Factor 5	-.10
Factor 6	-.14
Factor 7	-.14
Factor 8	-.20

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
instrlearn1	0.65	.006	0.64	0.67
instrlearn2	0.65	.007	0.63	0.66
instrlearn3	0.48	.008	0.47	0.50
instrlearn4	0.70	.006	0.69	0.71
replearn1	0.67	.006	0.66	0.68
replearn2	0.59	.007	0.58	0.61
replearn3	0.60	.007	0.59	0.62
replearn4	0.70	.006	0.69	0.71

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
instrlearn1	4.6	1.4	1	6	11031
instrlearn2	3.8	1.4	1	6	11001
instrlearn3	3.3	1.5	1	6	10993
instrlearn4	4.6	1.4	1	6	11052
replearn1	4.4	1.4	1	6	11041
replearn2	4.3	1.3	1	6	10990
replearn3	3.6	1.4	1	6	10991
replearn4	4.3	1.4	1	6	11010

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	4066	88	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	117	14	.000
Strong Invariance (Plus Equal Intercepts)	1511	14	.000
Strict Invariance (Plus Equal Error Variances)	337	14	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.994
French vs. Italian Language Version	.996
Italian vs. German Language Version	.990

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.999
Language: French	.998
Language: Italian	.993

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
instreplearn_fs	0.0	0.8	-2.7	1.5	11069

Share of Cases with Imputed Missing Values: 1.3%

(Equivalence of Scores from Robust MLMV: CD = .997)

Scale: Instructivist Learning: Teachers' Instructions

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	605	2	.000
	Baseline vs. Saturated	9077	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.165
	90% Confidence Interval: Lower Bound	.154
	90% Confidence Interval: Upper Bound	.176
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	147556
	Bayesian Information Criterion (BIC)	147643
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.934
	Tucker-Lewis Index (TLI)	.801
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.045
	Coefficient of Determination (CD)	.741

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.723
(Cronbach's alpha = .683)	
McDonald's Omega	.727

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.48
Factor 2	.05
Factor 3	-.12
Factor 4	-.22

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
instrlearn1	0.66	.008	0.65	0.68
instrlearn2	0.68	.008	0.67	0.70
instrlearn3	0.49	.009	0.47	0.51
instrlearn4	0.69	.008	0.67	0.70

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
instrlearn1	4.6	1.4	1	6	11031
instrlearn2	3.8	1.4	1	6	11001
instrlearn3	3.3	1.5	1	6	10993
instrlearn4	4.6	1.4	1	6	11052

Scale: Instructivist Learning: Teachers' Instructions (continued) Math Module

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	2118	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	49	6	.000
Strong Invariance (Plus Equal Intercepts)	466	6	.000
Strict Invariance (Plus Equal Error Variances)	146	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.994
French vs. Italian Language Version	.975
Italian vs. German Language Version	.978

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.998
Language: French	.998
Language: Italian	.958

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
instrlearn_fs	0.0	0.8	-2.6	1.4	11064

Share of Cases with Imputed Missing Values: 1.1%

(Equivalence of Scores from Robust MLMV: CD = .989)

Scale: Instructivist Learning: Repetitive Practice

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	24	2	.000
	Baseline vs. Saturated	9920	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.032
	90% Confidence Interval: Lower Bound	.021
	90% Confidence Interval: Upper Bound	.043
	Probability RMSEA <= 0.05	.996
3)	**Akaike's Information Criterion (AIC)**	145662
	Bayesian Information Criterion (BIC)	145750
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.998
	Tucker-Lewis Index (TLI)	.993
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.008
	Coefficient of Determination (CD)	.774

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.745
(Cronbach's alpha = .713)	
McDonald's Omega	.751

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.58
Factor 2	-.08
Factor 3	-.10
Factor 4	-.16

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
replearn1	0.76	.007	0.75	0.78
replearn2	0.71	.007	0.70	0.72
replearn3	0.49	.009	0.48	0.51
replearn4	0.64	.007	0.63	0.66

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
replearn1	4.4	1.4	1	6	11041
replearn2	4.3	1.3	1	6	10990
replearn3	3.6	1.4	1	6	10991
replearn4	4.3	1.4	1	6	11010

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1353	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	36	6	.000
Strong Invariance (Plus Equal Intercepts)	965	6	.000
Strict Invariance (Plus Equal Error Variances)	209	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.996
French vs. Italian Language Version	.999
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.999
Language: French	.996
Language: Italian	.997

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
replearn_fs	0.0	0.9	-2.8	1.5	11067

Share of Cases with Imputed Missing Values: 1.1%
(Equivalence of Scores from Robust MLMV: CD = .997)

Scale: System Aspect

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	2443	9	.000
	Baseline vs. Saturated	31459	15	.000

2)	**Root Mean Squared Error (RMSEA)**	.157
	90% Confidence Interval: Lower Bound	.152
	90% Confidence Interval: Upper Bound	.162
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	185422
	Bayesian Information Criterion (BIC)	185553
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.923
	Tucker-Lewis Index (TLI)	.871
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.050
	Coefficient of Determination (CD)	.879

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.878
(Cronbach's alpha = .854)	
McDonald's Omega	.878

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.21
Factor 2	.22
Factor 3	-.03
Factor 4	-.06
Factor 5	-.13
Factor 6	-.15

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
formasp1	0.71	.006	0.70	0.73
formasp2	0.72	.005	0.71	0.73
formasp3	0.75	.005	0.74	0.76
systasp1	0.74	.005	0.73	0.75
systasp2	0.76	.005	0.75	0.77
systasp3	0.75	.005	0.74	0.76

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
formasp1	4.3	1.3	1	6	10946
formasp2	4.1	1.3	1	6	10932
formasp3	4.4	1.2	1	6	10965
systasp1	5.0	1.2	1	6	10967
systasp2	4.7	1.2	1	6	10925
systasp3	4.7	1.2	1	6	10975

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	478	54	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	64	10	.000
Strong Invariance (Plus Equal Intercepts)	171	10	.000
Strict Invariance (Plus Equal Error Variances)	45	10	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.998
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.998

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
sysformasp_fs	0.0	0.8	-3.2	1.3	11006

Share of Cases with Imputed Missing Values: 1.3%
(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: System Aspect: Logical Thinking

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	12550	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			92905
	Bayesian Information Criterion (BIC)			92970
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.833

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.832
(Cronbach's alpha = .792)	
McDonald's Omega	.832

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.70
Factor 2	-.13
Factor 3	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
systasp1	0.76	.006	0.75	0.78
systasp2	0.81	.005	0.79	0.82
systasp3	0.80	.005	0.79	0.81

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
systasp1	5.0	1.2	1	6	10967
systasp2	4.7	1.2	1	6	10925
systasp3	4.7	1.2	1	6	10975

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	210	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	35	4	.000
Strong Invariance (Plus Equal Intercepts)	84	4	.000
Strict Invariance (Plus Equal Error Variances)	13	4	.012

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.997
French vs. Italian Language Version	1.000
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.997
Language: Italian	.995

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
systasp_fs	0.0	0.8	-3.1	1.0	11004

Share of Cases with Imputed Missing Values: 1.0%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: System Aspect: Formalism

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	11712	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			97123
	Bayesian Information Criterion (BIC)			97189
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.822

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.821
(Cronbach's alpha = .791)	
McDonald's Omega	.821

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.65
Factor 2	-.14
Factor 3	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
formasp1	0.78	.006	0.77	0.79
formasp2	0.79	.006	0.78	0.80
formasp3	0.77	.006	0.76	0.78

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
formasp1	4.3	1.3	1	6	10946
formasp2	4.1	1.3	1	6	10932
formasp3	4.4	1.2	1	6	10965

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	193	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	11	4	.025
Strong Invariance (Plus Equal Intercepts)	83	4	.000
Strict Invariance (Plus Equal Error Variances)	14	4	.008

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.994
Italian vs. German Language Version	.993

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.985

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
formasp_fs	0.0	0.9	-2.7	1.5	10992

Share of Cases with Imputed Missing Values: 0.9%
(Equivalence of Scores from Robust MLMV: CD = 1.00)

Scale: Scheme Aspect

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	12713	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			100471
	Bayesian Information Criterion (BIC)			100537
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.843

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.832
(Cronbach's alpha = .806)	
McDonald's Omega	.833

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.72
Factor 2	-.11
Factor 3	-.16

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
schemasp1	0.76	.006	0.75	0.77
schemasp2	0.76	.006	0.75	0.77
schemasp3	0.85	.005	0.84	0.86

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
schemasp1	3.9	1.4	1	6	10967
schemasp2	4.0	1.3	1	6	10926
schemasp3	3.7	1.4	1	6	10927

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	313	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	8	4	.092
Strong Invariance (Plus Equal Intercepts)	98	4	.000
Strict Invariance (Plus Equal Error Variances)	25	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	1.000
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.998

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
schemasp_fs	0.0	0.9	-2.4	1.8	10990

Share of Cases with Imputed Missing Values: 0.9%
(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Application Aspect

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	316	2	.000
	Baseline vs. Saturated	20302	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.119
	90% Confidence Interval: Lower Bound	.109
	90% Confidence Interval: Upper Bound	.131
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	129471
	Bayesian Information Criterion (BIC)	129559
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.985
	Tukker-Lewis Index (TLI)	.954
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.021
	Coefficient of Determination (CD)	.866

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.863
(Cronbach's alpha = .839)	
McDonald's Omega	.864

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.33
Factor 2	-.03
Factor 3	-.11
Factor 4	-.13

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
applyasp1	0.80	.005	0.79	0.81
applyasp2	0.79	.005	0.78	0.80
applyasp3	0.73	.005	0.72	0.74
applyasp4	0.81	.005	0.80	0.82

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
applyasp1	4.2	1.3	1	6	10982
applyasp2	4.6	1.3	1	6	10933
applyasp3	3.9	1.4	1	6	10958
applyasp4	4.3	1.3	1	6	10924

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	498	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	70	6	.000
Strong Invariance (Plus Equal Intercepts)	151	6	.000
Strict Invariance (Plus Equal Error Variances)	53	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.997
French vs. Italian Language Version	.992
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.999

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
applyasp_fs	0.0	1.0	-3.0	1.6	11007

Share of Cases with Imputed Missing Values: 1.1%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Father's Social Norms About Mathematics

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	12780	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			66659
	Bayesian Information Criterion (BIC)			66724
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.881

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.789
(Cronbach's alpha = .715)	
McDonald's Omega	.812

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.66
Factor 2	-.05
Factor 3	-.15

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
socnormsm1	0.87	.006	0.86	0.88
socnormsm2	0.89	.006	0.88	0.91
socnormsm3	0.50	.008	0.49	0.52

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
socnormsm1	3.2	0.7	1	4	10833
socnormsm2	3.1	0.8	1	4	10834
socnormsm3	2.4	0.9	1	4	10795

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
socnormsm1	3.95	-8.08	-4.66	1.62
socnormsm2	3.36	-5.95	-2.64	1.65
socnormsm3	0.99	-1.65	0.37	2.19

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	195	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	11	4	.030
Strong Invariance (Plus Equal Intercepts)	44	4	.000
Strict Invariance (Plus Equal Error Variances)	80	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.998
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.999
Language: French	.990
Language: Italian	.999

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
socnormsm_fs	0.1	0.9	-2.3	1.4	10847

Share of Cases with Imputed Missing Values: 0.6%
(Equivalence of Scores from Robust MLMV: CD = .996)
(Equivalence of Scores from Two-Step Approach: CD = .971)

Scale: Mother's Social Norms About Mathematics

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	15486	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			60431
	Bayesian Information Criterion (BIC)			60496
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.922

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.837
(Cronbach's alpha = .771)	
McDonald's Omega	.851

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.85
Factor 2	-.04
Factor 3	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
socnormsf1	0.95	.004	0.94	0.96
socnormsf2	0.85	.005	0.84	0.86
socnormsf3	0.60	.007	0.59	0.62

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
socnormsf1	3.3	0.7	1	4	10576
socnormsf2	3.2	0.8	1	4	10572
socnormsf3	3.1	0.9	1	4	10567

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
socnormsf1	4.84	-9.33	-5.83	1.21
socnormsf2	3.14	-5.97	-3.09	1.20
socnormsf3	1.25	-2.99	-1.28	0.85

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	198	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	15	4	.005
Strong Invariance (Plus Equal Intercepts)	85	4	.000
Strict Invariance (Plus Equal Error Variances)	72	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.999
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.999
Language: French	.996
Language: Italian	.956

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
socnormsf_fs	0.1	0.9	-2.4	1.2	10587

Share of Cases with Imputed Missing Values: 0.4%
(Equivalence of Scores from Robust MLMV: CD = .992)
(Equivalence of Scores from Two-Step Approach: CD = .960)

Scale: Achievement Expectations Parents

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	8040	2	.000
	Baseline vs. Saturated	24621	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.606
	90% Confidence Interval: Lower Bound	.595
	90% Confidence Interval: Upper Bound	.617
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	77644
	Bayesian Information Criterion (BIC)	77731
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.673
	Tucker-Lewis Index (TLI)	.020
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.108
	Coefficient of Determination (CD)	.854

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.837
(Cronbach's alpha = .774)	
McDonald's Omega	.834

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.35
Factor 2	.43
Factor 3	.11
Factor 4	-.19

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
expectf2	0.70	.007	0.69	0.72
expectf3	0.85	.005	0.84	0.86
expectm2	0.63	.009	0.62	0.65
expectm3	0.79	.005	0.78	0.80

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
expectf2	3.4	0.7	1	4	10568
expectf3	3.3	0.7	1	4	10566
expectm2	3.4	0.7	1	4	10862
expectm3	3.4	0.7	1	4	10864

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
expectf2	2.12	-5.87	-4.04	-0.32
expectf3	2.31	-5.88	-3.69	0.30
expectm2	1.75	-5.42	-3.28	0.14
expectm3	2.11	-6.40	-4.13	-0.12

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	297	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	15	6	.017
Strong Invariance (Plus Equal Intercepts)	126	6	.000
Strict Invariance (Plus Equal Error Variances)	12	6	.072

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.996
Italian vs. German Language Version	.995

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.964

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
expectp_fs	0.0	0.9	-3.1	1.1	10952

Share of Cases with Imputed Missing Values: 4.3%
(Equivalence of Scores from Robust MLMV: CD = .991)
(Equivalence of Scores from Two-Step Approach: CD = .941)

Scale: Achievement Expectations: Mother

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	4828	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			67851
	Bayesian Information Criterion (BIC)			67917
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.729

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.642
(Cronbach's alpha = .552)	
McDonald's Omega	.663

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.01
Factor 2	-.07
Factor 3	-.22

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
expectm1	0.42	.010	0.40	0.44
expectm2	0.80	.013	0.77	0.82
expectm3	0.65	.011	0.63	0.67

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
expectm1	2.8	0.8	1	4	10859
expectm2	3.4	0.7	1	4	10862
expectm3	3.4	0.7	1	4	10864

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
expectm1	0.83	-2.97	-0.79	1.48
expectm2	2.27	-6.07	-3.61	0.24
expectm3	1.68	-5.59	-3.50	-0.04

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	536	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	112	4	.000
Strong Invariance (Plus Equal Intercepts)	126	4	.000
Strict Invariance (Plus Equal Error Variances)	66	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.965
French vs. Italian Language Version	.982
Italian vs. German Language Version	.979

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.964
Language: French	.961
Language: Italian	.970

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
expectm_fs	0.0	0.8	-2.8	1.2	10864

Share of Cases with Imputed Missing Values: 0.1%
(Equivalence of Scores from Robust MLMV: CD = .987)
(Equivalence of Scores from Two-Step Approach: CD = .957)

Scale: Achievement Expectations: Father

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	7517	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			65854
	Bayesian Information Criterion (BIC)			65920
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.791

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.738
(Cronbach's alpha = .653)	
McDonald's Omega	.749

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.31
Factor 2	-.09
Factor 3	-.19

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
expectf1	0.55	.008	0.53	0.56
expectf2	0.83	.008	0.82	0.85
expectf3	0.72	.008	0.70	0.74

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
expectf1	2.9	0.9	1	4	10565
expectf2	3.4	0.7	1	4	10568
expectf3	3.3	0.7	1	4	10566

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
expectf1	1.17	-3.07	-1.05	1.32
expectf2	3.04	-7.28	-4.84	-0.32
expectf3	1.92	-5.13	-3.06	0.33

Scale: Achievement Expectations: Father (continued) Math Module

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	429	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	100	4	.000
Strong Invariance (Plus Equal Intercepts)	57	4	.000
Strict Invariance (Plus Equal Error Variances)	84	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.986
French vs. Italian Language Version	.997
Italian vs. German Language Version	.990

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.997
Language: French	.998
Language: Italian	.982

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
expectf_fs	0.0	0.8	-2.7	1.2	10569

Share of Cases with Imputed Missing Values: 0.1%

(Equivalence of Scores from Robust MLMV: CD = .988)

(Equivalence of Scores from Two-Step Approach: CD = .957)

Scale: Classmates' Appreciation of Mathematics

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	19804	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	53455
	Bayesian Information Criterion (BIC)	53521
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.946

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.834
(Cronbach's alpha = .776)	
McDonald's Omega	.859

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.94
Factor 2	-.02
Factor 3	-.08

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
apprmath1	0.92	.004	0.92	0.93
apprmath2	0.96	.004	0.95	0.97
apprmath3	0.53	.007	0.51	0.54

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
apprmath1	2.0	0.7	1	4	10778
apprmath2	2.0	0.7	1	4	10775
apprmath3	2.7	0.8	1	4	10776

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
apprmath1	4.34	-2.78	3.80	8.49
apprmath2	4.83	-2.94	4.63	9.65
apprmath3	1.14	-2.82	-0.55	2.41

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices *	chi2	df	p > chi2
	320	9	.000

Tests of Measurement Invariance *	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	13	2	.001
Strong Invariance (Plus Equal Intercepts)	67	2	.000
Strict Invariance (Plus Equal Error Variances)	5	2	.082

Configural Factor Similarity *

Tucker's Congruence Coefficient	TCC
German vs. French or Italian Version	1.000

Factor Score Equivalence: Group-Specific vs. Invariant Model *

Coefficient of Determination	CD
Language: German	.999
Language: French/ Italian	.991

* Note: Due to sparse tables for the italian version of the scale, equivalence tests failed to converge and were reestimated with collapsed italian and french versions.

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
apprmath_fs	0.0	0.9	-1.6	2.4	10784

Share of Cases with Imputed Missing Values: 0.1%
(Equivalence of Scores from Robust MLMV: CD = .997)
(Equivalence of Scores from Two-Step Approach: CD = .98)

Scale: Classroom Participation

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	584	5	.000
	Baseline vs. Saturated	28718	10	.000

2)	**Root Mean Squared Error (RMSEA)**	.103
	90% Confidence Interval: Lower Bound	.096
	90% Confidence Interval: Upper Bound	.110
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	97128
	Bayesian Information Criterion (BIC)	97238
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.980
	Tucker-Lewis Index (TLI)	.960
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.024
	Coefficient of Determination (CD)	.890

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.888
(Cronbach's alpha = .848)	
McDonald's Omega	.888

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.95
Factor 2	.02
Factor 3	-.05
Factor 4	-.11
Factor 5	-.11

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
engage1	0.76	.005	0.75	0.77
engage2	0.83	.004	0.82	0.84
engage3	0.75	.005	0.74	0.76
engage4	0.80	.004	0.79	0.81
engage5	0.77	.005	0.76	0.78

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
engage1	2.9	0.8	1	4	10897
engage2	2.9	0.7	1	4	10852
engage3	3.0	0.7	1	4	10907
engage4	3.0	0.8	1	4	10898
engage5	2.8	0.8	1	4	10829

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
engage1	2.22	-4.53	-1.82	2.06
engage2	2.82	-5.44	-2.01	3.03
engage3	2.14	-4.97	-2.11	1.89
engage4	2.51	-5.30	-2.40	2.21
engage5	2.28	-4.28	-1.30	3.10

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	938	40	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	51	8	.000
Strong Invariance (Plus Equal Intercepts)	31	8	.000
Strict Invariance (Plus Equal Error Variances)	149	8	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.997
French vs. Italian Language Version	.997
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.999

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
engage_fs	0.0	0.9	-2.7	1.9	10936

Share of Cases with Imputed Missing Values: 1.5%
(Equivalence of Scores from Robust MLMV: CD = .996)
(Equivalence of Scores from Two-Step Approach: CD = .984)

Scale: Identified Motivation (Mathematics) Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	45	2	.000
	Baseline vs. Saturated	43936	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.044
	90% Confidence Interval: Lower Bound	.034
	90% Confidence Interval: Upper Bound	.056
	Probability RMSEA <= 0.05	.777
3)	**Akaike's Information Criterion (AIC)**	72033
	Bayesian Information Criterion (BIC)	72121
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.999
	Tucker-Lewis Index (TLI)	.997
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.004
	Coefficient of Determination (CD)	.955

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.946
(Cronbach's alpha = .918)	
McDonald's Omega	.947

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.20
Factor 2	-.04
Factor 3	-.05
Factor 4	-.04

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
instrumot1	0.95	.001	0.94	0.95
instrumot2	0.93	.002	0.93	0.94
instrumot3	0.89	.002	0.88	0.89
instrumot4	0.85	.003	0.84	0.85

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
instrumot1	2.9	0.9	1	4	11018
instrumot2	2.9	0.9	1	4	11020
instrumot3	2.8	0.9	1	4	11030
instrumot4	2.9	0.9	1	4	11013

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators *	Coef.	Cut1	Cut2	Cut3
instrumot1	4.16	-7.00	-2.77	2.59
instrumot2	3.66	-5.86	-2.07	1.94
instrumot3	2.86	-5.38	-1.92	2.16
instrumot4	2.49	-5.04	-2.19	1.86

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices *	chi2	df	p > chi2
	387	14	.000

Tests of Measurement Invariance *	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	111	3	.000
Strong Invariance (Plus Equal Intercepts)	75	3	.000
Strict Invariance (Plus Equal Error Variances)	135	3	.000

Configural Factor Similarity *	
Tucker's Congruence Coefficient	TCC
German vs. French or Italian Version	1.000

Factor Score Equivalence: Group-Specific vs. Invariant Model *	
Coefficient of Determination	CD
Language: German	1.000
Language: French/ Italian	1.000

* Note: Due to sparse tables for the italian version of the scale, equivalence tests failed to converge and were reestimated with collapsed italian and french versions.

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
instrumot_fs	-0.1	1.0	-2.4	1.5	11033

Share of Cases with Imputed Missing Values: 0.3%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .985)

Scale: External Motivation Regulation

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	687	2	.000
	Baseline vs. Saturated	16452	6	.000
2)	**Root Mean Squared Error (RMSEA)**			.177
	90% Confidence Interval: Lower Bound			.166
	90% Confidence Interval: Upper Bound			.188
	Probability RMSEA <= 0.05			.000
3)	**Akaike's Information Criterion (AIC)**			100910
	Bayesian Information Criterion (BIC)			100998
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			.958
	Tucker-Lewis Index (TLI)			.875
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.038
	Coefficient of Determination (CD)			.844

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.820
(Cronbach's alpha = .764)	
McDonald's Omega	.826

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.06
Factor 2	.06
Factor 3	-.15
Factor 4	-.15

Standardized Factor Loadings

Indicators *	Coef.	(SE)	[95% Conf.	Interval]
extreg2	0.76	.005	0.75	0.77
extreg3	0.81	.005	0.80	0.82
extreg4	0.58	.008	0.56	0.59
extreg5	0.78	.005	0.77	0.79

*** Note:** Items Extreg1 and Extreg6 Excluded to Improve Scale Quality

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
extreg2	1.9	0.9	1	4	10901
extreg3	2.0	0.9	1	4	10830
extreg4	2.4	0.9	1	4	10841
extreg5	1.8	0.9	1	4	10827

*** Note:** Items Extreg1 and Extreg6 Excluded to Improve Scale Quality

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators *	Coef.	Cut1	Cut2	Cut3
extreg2	2.11	-0.76	1.62	4.25
extreg3	2.55	-1.03	1.52	4.56
extreg4	1.28	-1.75	0.01	2.39
extreg5	2.34	-0.17	2.28	4.99

*** Note:** Items Extreg1 and Extreg6 Excluded to Improve Scale Quality

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	222	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	46	6	.000
Strong Invariance (Plus Equal Intercepts)	113	6	.000
Strict Invariance (Plus Equal Error Variances)	35	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.990
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	.997

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
extregm_fs	0.0	0.9	-1.4	2.5	10930

Share of Cases with Imputed Missing Values: 1.5%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .977)

Scale: Dispositional Interest

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	1805	9	.000
	Baseline vs. Saturated	31076	15	.000

2)	**Root Mean Squared Error (RMSEA)**	.135
	90% Confidence Interval: Lower Bound	.130
	90% Confidence Interval: Upper Bound	.140
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	137195
	Bayesian Information Criterion (BIC)	137326
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.942
	Tucker-Lewis Index (TLI)	.904
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.041
	Coefficient of Determination (CD)	.888

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.875
(Cronbach's alpha = .836)	
McDonald's Omega	.876

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.19
Factor 2	.14
Factor 3	-.01
Factor 4	-.05
Factor 5	-.13
Factor 6	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
intsubj1	0.84	.004	0.83	0.85
intsubj2	0.65	.006	0.64	0.66
intsubj3	0.75	.005	0.74	0.76
intsubj4	0.66	.006	0.65	0.67
intsubj5	0.69	.006	0.68	0.71
intsubj6	0.80	.004	0.80	0.81

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
intsubj1	2.5	0.9	1	4	10889
intsubj2	3.2	0.7	1	4	10922
intsubj3	2.9	0.8	1	4	10845
intsubj4	2.6	0.9	1	4	10842
intsubj5	2.8	0.8	1	4	10905
intsubj6	2.4	1.0	1	4	10853

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
intsubj1	2.92	-3.37	-0.29	3.76
intsubj2	1.58	-4.54	-2.81	0.59
intsubj3	2.12	-4.06	-1.70	1.90
intsubj4	1.63	-2.34	-0.39	2.29
intsubj5	1.80	-3.88	-0.89	2.43
intsubj6	2.53	-2.10	0.31	3.26

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	885	54	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	91	10	.000
Strong Invariance (Plus Equal Intercepts)	332	10	.000
Strict Invariance (Plus Equal Error Variances)	77	10	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.996
French vs. Italian Language Version	.995
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.999

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
intsubj_fs	0.0	0.9	-2.6	2.1	10949

Share of Cases with Imputed Missing Values: 1.6%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .988)

Scale: Performance-Approach Goals (SELLMO) Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	620	2	.000
	Baseline vs. Saturated	17637	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.171
	90% Confidence Interval: Lower Bound	.159
	90% Confidence Interval: Upper Bound	.182
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	117025
	Bayesian Information Criterion (BIC)	117112
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.965
	Tucker-Lewis Index (TLI)	.895
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.040
	Coefficient of Determination (CD)	.865

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.834
(Cronbach's alpha = .804)	
McDonald's Omega	.837

Test of (One-)Dimensionality (Parallel Analysis)
Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.16
Factor 2	.05
Factor 3	-.15
Factor 4	-.13

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
approxgoals1	0.74	.006	0.73	0.75
approxgoals2	0.84	.004	0.83	0.84
approxgoals3	0.57	.008	0.55	0.58
approxgoals4	0.84	.004	0.83	0.85

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
approxgoals1	2.8	1.2	1	5	10608
approxgoals2	2.5	1.2	1	5	10478
approxgoals3	3.3	1.1	1	5	10596
approxgoals4	2.7	1.2	1	5	10474

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	370	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	51	6	.000
Strong Invariance (Plus Equal Intercepts)	89	6	.000
Strict Invariance (Plus Equal Error Variances)	76	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.988
Italian vs. German Language Version	.985

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.991

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
approxgoals_fs	0.0	0.8	-1.4	1.9	10628

Share of Cases with Imputed Missing Values: 1.8%
(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Learning Goal Orientation (SELLMO) Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	396	2	.000
	Baseline vs. Saturated	16559	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.136
	90% Confidence Interval: Lower Bound	.125
	90% Confidence Interval: Upper Bound	.147
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	113590
	Bayesian Information Criterion (BIC)	113677
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.976
	Tucker-Lewis Index (TLI)	.929
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.028
	Coefficient of Determination (CD)	.841

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.839
(Cronbach's alpha = .808)	
McDonald's Omega	.839

Test of (One-)Dimensionality (Parallel Analysis)
Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.15
Factor 2	-.01
Factor 3	-.15
Factor 4	-.13

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
learntarget1	0.74	.006	0.72	0.75
learntarget2	0.76	.006	0.75	0.77
learntarget3	0.73	.006	0.72	0.74
learntarget4	0.78	.005	0.77	0.79

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
learntarget1	3.3	1.1	1	5	10637
learntarget2	3.4	1.1	1	5	10481
learntarget3	3.3	1.1	1	5	10606
learntarget4	3.1	1.1	1	5	10485

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	887	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	12	6	.072
Strong Invariance (Plus Equal Intercepts)	421	6	.000
Strict Invariance (Plus Equal Error Variances)	254	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.999
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.997

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
learntarget_fs	0.0	0.7	-2.0	1.5	10649

Share of Cases with Imputed Missing Values: 1.8%

(Equivalence of Scores from Robust MLMV: CD = .998)

Scale: Work Avoidance (SELLMO) Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	370	2	.000
	Baseline vs. Saturated	9625	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.131
	90% Confidence Interval: Lower Bound	.120
	90% Confidence Interval: Upper Bound	.143
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	122140
	Bayesian Information Criterion (BIC)	122227
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.962
	Tucker-Lewis Index (TLI)	.885
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.033
	Coefficient of Determination (CD)	.761

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.747
(Cronbach's alpha = .712)	
McDonald's Omega	.750

Test of (One-)Dimensionality (Parallel Analysis)
Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.59
Factor 2	-.02
Factor 3	-.09
Factor 4	-.22

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
avoidwork1	0.53	.009	0.51	0.54
avoidwork2	0.70	.007	0.68	0.71
avoidwork3	0.67	.008	0.66	0.69
avoidwork4	0.71	.007	0.70	0.72

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
avoidwork1	2.9	1.1	1	5	10615
avoidwork2	3.1	1.1	1	5	10483
avoidwork3	3.2	1.2	1	5	10599
avoidwork4	3.1	1.1	1	5	10480

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	611	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	11	6	.087
Strong Invariance (Plus Equal Intercepts)	282	6	.000
Strict Invariance (Plus Equal Error Variances)	170	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.989
Italian vs. German Language Version	.994

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	.991

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
avoidwork_fs	0.0	0.5	-1.2	1.2	10637

Share of Cases with Imputed Missing Values: 1.8%
(Equivalence of Scores from Robust MLMV: CD = .996)

Scale: Avoidance Performance Goals (SELLMO) Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	550	2	.000
	Baseline vs. Saturated	20651	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.160
	90% Confidence Interval: Lower Bound	.149
	90% Confidence Interval: Upper Bound	.172
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	117023
	Bayesian Information Criterion (BIC)	117111
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.973
	Tucker-Lewis Index (TLI)	.920
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.027
	Coefficient of Determination (CD)	.877

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.866
(Cronbach's alpha = .830)	
McDonald's Omega	.867

Test of (One-)Dimensionality (Parallel Analysis)
Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.37
Factor 2	.01
Factor 3	-.09
Factor 4	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
avoidblame1	0.73	.005	0.72	0.74
avoidblame2	0.75	.005	0.74	0.76
avoidblame3	0.86	.004	0.85	0.87
avoidblame4	0.81	.005	0.80	0.81

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
avoidblame1	2.6	1.2	1	5	10594
avoidblame2	2.6	1.3	1	5	10496
avoidblame3	2.5	1.2	1	5	10604
avoidblame4	2.3	1.1	1	5	10509

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	378	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	19	6	.004
Strong Invariance (Plus Equal Intercepts)	120	6	.000
Strict Invariance (Plus Equal Error Variances)	161	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.997
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
avoidblame_fs	0.0	0.8	-1.2	2.1	10642

Share of Cases with Imputed Missing Values: 1.9%
(Equivalence of Scores from Robust MLMV: CD = .998)

Scale: Self-Concept in Mathematics

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	57824	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	134733
	Bayesian Information Criterion (BIC)	134805
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.980

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.927
(Cronbach's alpha = .888)	
McDonald's Omega	.930

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.38
Factor 2	-.01
Factor 3	-.08

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
matcon1	0.90	.002	0.90	0.90
matcon2	0.99	.001	0.99	0.99
matcon3	0.82	.002	0.81	0.82

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
matcon1	2.7	0.9	1	4	22183
matcon2	2.4	1.1	1	4	22187
matcon3	2.4	1.0	1	4	22180

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
matcon1	3.38	-4.50	-1.06	2.95
matcon2	4.96	-3.20	0.21	4.25
matcon3	2.40	-2.30	0.21	2.53

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	937	18	.000

Tests of measurement invariance *	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	335	4	.000
Strong Invariance (Plus Equal Intercepts)	47	4	.000
Strict Invariance (Plus Equal Error Variances)	241	2	.000

Configural Factor Similary *	
Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.997
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models *	
Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	1.000

*** Note:** Model-bases invariance tests do not converge unless the error variance of item matcon2 is constrained to zero.

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
matcon_fs	0.0	1.0	-1.7	1.6	22193

Share of Cases with Imputed Missing Values: 0.1%
(Equivalence of Scores from Robust MLMV: CD = .967)
(Equivalence of Scores from Two-Step Approach: CD = .899)

Scale: Specific Self-Efficacy: Numeracy

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	536	2	.000
	Baseline vs. Saturated	36814	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.110
	90% Confidence Interval: Lower Bound	.103
	90% Confidence Interval: Upper Bound	.118
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	196455
	Bayesian Information Criterion (BIC)	196551
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.985
	Tucker-Lewis Index (TLI)	.956
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.020
	Coefficient of Determination (CD)	.854

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.851
(Cronbach's alpha = .831)	
McDonald's Omega	.852

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.23
Factor 2	-.05
Factor 3	-.08
Factor 4	-.16

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
selfeff01	0.77	.004	0.76	0.77
selfeff02	0.77	.004	0.76	0.78
selfeff03	0.80	.003	0.79	0.81
selfeff04	0.73	.004	0.72	0.74

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
selfeff01	3.3	0.9	1	4	21801
selfeff02	3.0	0.9	1	4	21827
selfeff03	2.8	0.9	1	4	10734
selfeff04	2.7	0.9	1	4	10755

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
selfeff01	2.35	-4.76	-2.62	-0.16
selfeff02	2.38	-4.13	-1.77	1.07
selfeff03	3.03	-5.40	-1.83	2.94
selfeff04	2.27	-4.13	-1.09	2.49

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	651	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	72	6	.000
Strong Invariance (Plus Equal Intercepts)	85	6	.000
Strict Invariance (Plus Equal Error Variances)	33	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.998
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
selfeffa_fs	0.0	0.9	-2.4	1.6	21881

Share of Cases with Imputed Missing Values: 51.2%
(Equivalence of Scores from Robust MLMV: CD = .995)
(Equivalence of Scores from Two-Step Approach: CD = .976)

Scale: Specific Self-Efficacy: Algebra

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	3889	2	.000
	Baseline vs. Saturated	92426	6	.000
2)	**Root Mean Squared Error (RMSEA)**			.298
	90% Confidence Interval: Lower Bound			.290
	90% Confidence Interval: Upper Bound			.306
	Probability RMSEA <= 0.05			.000
3)	**Akaike's Information Criterion (AIC)**			147967
	Bayesian Information Criterion (BIC)			148063
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			.958
	Tucker-Lewis Index (TLI)			.874
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.026
	Coefficient of Determination (CD)			.957

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.947
(Cronbach's alpha = .926)	
McDonald's Omega	.948

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.24
Factor 2	.07
Factor 3	-.06
Factor 4	-.06

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
selfeff05	0.86	.002	0.86	0.87
selfeff06	0.95	.001	0.95	0.96
selfeff07	0.88	.002	0.88	0.89
selfeff08	0.92	.001	0.92	0.93

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
selfeff05	3.3	0.9	1	4	21809
selfeff06	3.0	1.0	1	4	21794
selfeff07	2.8	1.0	1	4	10747
selfeff08	3.2	0.9	1	4	10730

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
selfeff05	3.39	-5.99	-3.58	-0.95
selfeff06	8.35	-11.55	-5.35	1.58
selfeff07	4.65	-6.43	-2.51	1.99
selfeff08	5.99	-9.89	-5.56	-0.57

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	506	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	17	6	.010
Strong Invariance (Plus Equal Intercepts)	116	6	.000
Strict Invariance (Plus Equal Error Variances)	238	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	1.000
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.998

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
selfeffb_fs	-0.1	0.9	-2.2	1.1	21872

Share of Cases with Imputed Missing Values: 51.2%
(Equivalence of Scores from Robust MLMV: CD = .998)
(Equivalence of Scores from Two-Step Approach: CD = .957)

Scale: Specific Self-Efficacy: Geometry

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	229	2	.000
	Baseline vs. Saturated	30977	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.072
	90% Confidence Interval: Lower Bound	.064
	90% Confidence Interval: Upper Bound	.080
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	203347
	Bayesian Information Criterion (BIC)	203443
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.993
	Tucker-Lewis Index (TLI)	.978
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.015
	Coefficient of Determination (CD)	.836

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.823
(Cronbach's alpha = .803)	
McDonald's Omega	.825

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.05
Factor 2	-.07
Factor 3	-.09
Factor 4	-.16

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
selfeff09	0.81	.004	0.80	0.81
selfeff10	0.76	.004	0.75	0.76
selfeff11	0.75	.004	0.74	0.75
selfeff12	0.63	.005	0.62	0.64

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
selfeff09	3.3	0.9	1	4	10752
selfeff10	3.2	0.9	1	4	21783
selfeff11	3.0	1.0	1	4	21802
selfeff12	2.6	0.9	1	4	10751

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
selfeff09	3.22	-6.78	-3.69	-0.03
selfeff10	2.24	-4.55	-2.29	0.17
selfeff11	2.15	-3.88	-1.49	0.85
selfeff12	1.75	-3.32	-0.62	2.77

Scale: Specific Self-Efficacy: Geometry (continued) Core Questionnaire

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	3499	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	59	6	.000
Strong Invariance (Plus Equal Intercepts)	2400	6	.000
Strict Invariance (Plus Equal Error Variances)	320	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.997
Italian vs. German Language Version	.993

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.993
Language: Italian	.988

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
selfeffc_fs	0.0	0.9	-2.5	1.5	21875

Share of Cases with Imputed Missing Values: 51.3%
(Equivalence of Scores from Robust MLMV: CD = .995)
(Equivalence of Scores from Two-Step Approach: CD = .965)

Scale: Specific Self-Efficacy: Probability

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	1326	2	.000
	Baseline vs. Saturated	63299	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.174
	90% Confidence Interval: Lower Bound	.166
	90% Confidence Interval: Upper Bound	.182
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	178726
	Bayesian Information Criterion (BIC)	178821
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.979
	Tucker-Lewis Index (TLI)	.937
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.022
	Coefficient of Determination (CD)	.919

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.917
(Cronbach's alpha = .907)	
McDonald's Omega	.917

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.86
Factor 2	.01
Factor 3	-.09
Factor 4	-.10

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
selfeff13	0.87	.002	0.86	0.87
selfeff14	0.84	.002	0.83	0.84
selfeff15	0.89	.002	0.88	0.89
selfeff16	0.83	.003	0.83	0.84

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
selfeff13	2.7	1.0	1	4	21778
selfeff14	2.6	1.0	1	4	10754
selfeff15	2.8	0.9	1	4	21776
selfeff16	2.5	0.9	1	4	10751

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
selfeff13	3.46	-4.44	-0.85	2.41
selfeff14	3.65	-4.88	-0.67	3.58
selfeff15	3.96	-5.27	-1.24	2.74
selfeff16	3.51	-4.69	-0.45	3.96

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	118	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	11	6	.102
Strong Invariance (Plus Equal Intercepts)	42	6	.000
Strict Invariance (Plus Equal Error Variances)	21	6	.002

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	1.000
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
selfeffd_fs	0.0	0.9	-2.0	1.7	21858

Share of Cases with Imputed Missing Values: 51.2%
(Equivalence of Scores from Robust MLMV: CD = .997)
(Equivalence of Scores from Two-Step Approach: CD = .986)

Scale: Mathematics Anxiety

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	1904	5	.000
	Baseline vs. Saturated	37885	10	.000

2)	**Root Mean Squared Error (RMSEA)**	.186
	90% Confidence Interval: Lower Bound	.179
	90% Confidence Interval: Upper Bound	.193
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	114426
	Bayesian Information Criterion (BIC)	114535
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.950
	Tucker-Lewis Index (TLI)	.900
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.035
	Coefficient of Determination (CD)	.916

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.914
(Cronbach's alpha = .877)	
McDonald's Omega	.914

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.35
Factor 2	.10
Factor 3	-.03
Factor 4	-.10
Factor 5	-.12

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
anxmath1	0.83	.004	0.82	0.84
anxmath2	0.79	.004	0.79	0.80
anxmath3	0.84	.004	0.83	0.85
anxmath4	0.80	.004	0.79	0.81
anxmath5	0.86	.003	0.85	0.86

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
anxmath1	2.4	1.0	1	4	10999
anxmath2	1.9	0.9	1	4	10996
anxmath3	1.8	0.9	1	4	10992
anxmath4	2.5	1.0	1	4	10995
anxmath5	2.1	1.0	1	4	10994

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
anxmath1	2.83	-2.61	0.30	3.40
anxmath2	2.48	-0.85	1.87	4.52
anxmath3	2.94	-0.26	2.70	5.39
anxmath4	2.50	-2.49	-0.24	2.32
anxmath5	3.11	-1.59	1.60	4.41

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1137	40	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	141	8	.000
Strong Invariance (Plus Equal Intercepts)	502	8	.000
Strict Invariance (Plus Equal Error Variances)	151	8	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.995
Italian vs. German Language Version	.988

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.980

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
anxmath_fs	0.0	0.9	-1.6	2.3	11005

Share of Cases with Imputed Missing Values: 0.2%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .976)

Scale: Mathematics Boredom

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	689	2	.000
	Baseline vs. Saturated	20215	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.178
	90% Confidence Interval: Lower Bound	.167
	90% Confidence Interval: Upper Bound	.189
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	125128
	Bayesian Information Criterion (BIC)	125216
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.966
	Tucker-Lewis Index (TLI)	.898
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.032
	Coefficient of Determination (CD)	.863

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.863
(Cronbach's alpha = .831)	
McDonald's Omega	.863

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.34
Factor 2	.02
Factor 3	-.11
Factor 4	-.15

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
boredom1	0.78	.005	0.77	0.79
boredom2	0.78	.005	0.77	0.79
boredom3	0.80	.005	0.79	0.81
boredom4	0.77	.005	0.76	0.78

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
boredom1	2.9	1.3	1	5	10877
boredom2	2.6	1.2	1	5	10834
boredom3	2.5	1.3	1	5	10813
boredom4	3.0	1.3	1	5	10877

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	815	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	15	6	.022
Strong Invariance (Plus Equal Intercepts)	599	6	.000
Strict Invariance (Plus Equal Error Variances)	166	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.997
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.995

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
boredom_fs	0.0	0.9	-1.5	1.9	10902

Share of Cases with Imputed Missing Values: 1.1%

(Equivalence of Scores from Robust MLMV: CD = .998)

Scale: Mathematics Anger

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	79	2	.000
	Baseline vs. Saturated	27251	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.059
	90% Confidence Interval: Lower Bound	.049
	90% Confidence Interval: Upper Bound	.071
	Probability RMSEA <= 0.05	.073
3)	**Akaike's Information Criterion (AIC)**	120644
	Bayesian Information Criterion (BIC)	120732
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.997
	Tucker-Lewis Index (TLI)	.992
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.010
	Coefficient of Determination (CD)	.915

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.895
(Cronbach's alpha = .865)	
McDonald's Omega	.897

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.66
Factor 2	-.05
Factor 3	-.08
Factor 4	-.09

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
anger1	0.70	.005	0.69	0.71
anger2	0.89	.003	0.89	0.90
anger3	0.89	.003	0.88	0.89
anger4	0.82	.004	0.82	0.83

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
anger1	2.6	1.2	1	5	10891
anger2	2.4	1.3	1	5	10815
anger3	2.5	1.3	1	5	10810
anger4	2.5	1.4	1	5	10869

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1045	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	52	6	.000
Strong Invariance (Plus Equal Intercepts)	264	6	.000
Strict Invariance (Plus Equal Error Variances)	48	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.997
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.996

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
anger_fs	0.0	0.9	-1.4	2.1	10902

Share of Cases with Imputed Missing Values: 1.1%
(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Mathematics Enjoyment

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	191	2	.000
	Baseline vs. Saturated	23069	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.093
	90% Confidence Interval: Lower Bound	.082
	90% Confidence Interval: Upper Bound	.104
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	114281
	Bayesian Information Criterion (BIC)	114369
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.992
	Tucker-Lewis Index (TLI)	.975
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.014
	Coefficient of Determination (CD)	.892

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.877
(Cronbach's alpha = .845)	
McDonald's Omega	.879

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.47
Factor 2	-.04
Factor 3	-.09
Factor 4	-.11

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
enjoymath1	0.86	.004	0.86	0.87
enjoymath2	0.86	.004	0.86	0.87
enjoymath3	0.73	.005	0.72	0.74
enjoymath4	0.75	.005	0.74	0.76

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
enjoymath1	2.5	1.2	1	5	10880
enjoymath2	2.5	1.2	1	5	10830
enjoymath3	2.3	1.2	1	5	10882
enjoymath4	2.3	1.1	1	5	10823

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	333	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	44	6	.000
Strong Invariance (Plus Equal Intercepts)	152	6	.000
Strict Invariance (Plus Equal Error Variances)	40	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	1.000
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.998

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
enjoymath_fs	0.0	0.9	-1.4	2.5	10907

Share of Cases with Imputed Missing Values: 1.0%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Positive Attitudes towards School

Core Questionnaire

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	22788	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			205667
	Bayesian Information Criterion (BIC)			205739
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.835

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.809
(Cronbach's alpha = .784)	
McDonald's Omega	.813

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.61
Factor 2	-.10
Factor 3	-.17

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
posatt1	0.74	.004	0.73	0.75
posatt2	0.86	.004	0.85	0.87
posatt3	0.70	.004	0.69	0.71

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
posatt1	3.8	1.3	1	6	22295
posatt2	4.1	1.3	1	6	22288
posatt3	4.6	1.3	1	6	22287

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	998	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	17	4	.002
Strong Invariance (Plus Equal Intercepts)	172	4	.000
Strict Invariance (Plus Equal Error Variances)	217	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	1.000
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
posatt_fs	0.0	0.9	-2.5	1.4	22299

Share of Cases with Imputed Missing Values: 0.1%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Enjoyment in School

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	24844	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			216963
	Bayesian Information Criterion (BIC)			217035
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.856

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.821
(Cronbach's alpha = .796)	
McDonald's Omega	.825

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.67
Factor 2	-.08
Factor 3	-.16

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
enjoyschool1	0.76	.004	0.75	0.77
enjoyschool2	0.89	.004	0.88	0.89
enjoyschool3	0.69	.004	0.68	0.70

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
enjoyschool1	3.2	1.5	1	6	22254
enjoyschool2	3.5	1.4	1	6	22252
enjoyschool3	3.9	1.4	1	6	22257

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	506	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	33	4	.000
Strong Invariance (Plus Equal Intercepts)	258	4	.000
Strict Invariance (Plus Equal Error Variances)	34	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.992
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	.994

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
enjoyschool_fs	0.0	1.1	-2.1	2.1	22267

Share of Cases with Imputed Missing Values: 0.1%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Academic Self-Efficacy

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	32752	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	179405
	Bayesian Information Criterion (BIC)	179477
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.874

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.868
(Cronbach's alpha = .836)	
McDonald's Omega	.869

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.92
Factor 2	-.11
Factor 3	-.13

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
acaself1	0.81	.003	0.80	0.81
acaself2	0.87	.003	0.87	0.88
acaself3	0.81	.003	0.80	0.81

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
acaself1	4.7	1.1	1	6	22256
acaself2	4.1	1.2	1	6	22248
acaself3	4.3	1.2	1	6	22252

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	774	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	77	4	.000
Strong Invariance (Plus Equal Intercepts)	250	4	.000
Strict Invariance (Plus Equal Error Variances)	318	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.998
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.989

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
acaself_fs	0.0	0.8	-2.7	1.4	22264

Share of Cases with Imputed Missing Values: 0.1%
(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Physical Complaints in School

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	29	2	.000
	Baseline vs. Saturated	36796	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.025
	90% Confidence Interval: Lower Bound	.017
	90% Confidence Interval: Upper Bound	.033
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	272002
	Bayesian Information Criterion (BIC)	272098
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.999
	Tucker-Lewis Index (TLI)	.998
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.005
	Coefficient of Determination (CD)	.857

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.847
(Cronbach's alpha = .772)	
McDonald's Omega	.849

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.22
Factor 2	-.09
Factor 3	-.10
Factor 4	-.12

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
physpain1	0.78	.003	0.77	0.79
physpain2	0.79	.003	0.78	0.79
physpain3	0.82	.003	0.81	0.82
physpain4	0.67	.004	0.66	0.68

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
physpain1	1.7	1.3	1	6	22260
physpain2	1.7	1.4	1	6	22249
physpain3	1.7	1.3	1	6	22222
physpain4	2.3	1.6	1	6	22245

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1179	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	76	6	.000
Strong Invariance (Plus Equal Intercepts)	188	6	.000
Strict Invariance (Plus Equal Error Variances)	542	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.997
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.988

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
physpain_fs	0.0	0.8	-.6	3.5	22271

Share of Cases with Imputed Missing Values: 0.3%
(Equivalence of Scores from Robust MLMV: CD = .995)

Scale: Worries about School

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	21848	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	240309
	Bayesian Information Criterion (BIC)	240381
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.836

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.795
(Cronbach's alpha = .753)	
McDonald's Omega	.802

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.57
Factor 2	-.09
Factor 3	-.18

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
trouschool1	0.78	.004	0.78	0.79
trouschool2	0.86	.004	0.85	0.87
trouschool3	0.62	.005	0.61	0.63

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
trouschool1	2.9	1.6	1	6	22260
trouschool2	3.2	1.7	1	6	22263
trouschool3	3.4	1.9	1	6	22263

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1522	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	51	4	.000
Strong Invariance (Plus Equal Intercepts)	889	4	.000
Strict Invariance (Plus Equal Error Variances)	295	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.999
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.996

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
trouschool_fs	0.0	1.2	-1.9	2.5	22270

Share of Cases with Imputed Missing Values: 0.1%

(Equivalence of Scores from Robust MLMV: CD = .997)

Scale: Social Problems in School

Core Questionnaire

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	39687	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			164458
	Bayesian Information Criterion (BIC)			164530
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.929

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.886
(Cronbach's alpha = .817)	
McDonald's Omega	.889

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.07
Factor 2	-.05
Factor 3	-.12

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
socprob1	0.95	.002	0.95	0.95
socprob2	0.84	.003	0.84	0.85
socprob3	0.76	.003	0.75	0.77

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
socprob1	1.5	1.0	1	6	22244
socprob2	1.7	1.2	1	6	22259
socprob3	1.5	1.1	1	6	22239

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	466	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	16	4	.003
Strong Invariance (Plus Equal Intercepts)	129	4	.000
Strict Invariance (Plus Equal Error Variances)	157	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.999
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
socprob_fs	0.0	0.9	-0.5	4.3	22265

Share of Cases with Imputed Missing Values: 0.2%

(Equivalence of Scores from Robust MLMV: CD = .991)

Scale: School Reluctance

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	14239	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	245338
	Bayesian Information Criterion (BIC)	245410
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.835

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.702
(Cronbach's alpha = .661)	
McDonald's Omega	.727

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.23
Factor 2	-.05
Factor 3	-.22

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
schoolav1	0.89	.007	0.88	0.91
schoolav2	0.67	.007	0.66	0.69
schoolav3	0.46	.006	0.45	0.47

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
schoolav1	3.1	1.8	1	6	22245
schoolav2	3.7	1.9	1	6	22248
schoolav3	2.2	1.5	1	6	22235

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices *	chi2	df	p > chi2
	1451	9	.000

Tests of Measurement Invariance *	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	99	2	.000
Strong Invariance (Plus Equal Intercepts)	981	2	.000
Strict Invariance (Plus Equal Error Variances)	49	2	.000

Configural Factor Similarity *	
Tucker's Congruence Coefficient	TCC
German vs. French or Italian Version	.999

Factor Score Equivalence: Group-Specific vs. Invariant Model *	
Coefficient of Determination	CD
Language: German	.994
Language: French/ Italian	.981

* Note: Due to sparse tables for the italian version of the scale, equivalence tests failed to converge and were reestimated with collapsed italian and french versions.

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
schoolav_fs	0.0	1.4	-2.0	2.6	22266

Share of Cases with Imputed Missing Values: 0.2%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Teacher: Cognitive Activation

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	5636	20	.000
	Baseline vs. Saturated	38613	28	.000

2)	**Root Mean Squared Error (RMSEA)**	.164
	90% Confidence Interval: Lower Bound	.160
	90% Confidence Interval: Upper Bound	.167
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	176245
	Bayesian Information Criterion (BIC)	176419
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.854
	Tucker-Lewis Index (TLI)	.796
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.078
	Coefficient of Determination (CD)	.894

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.873
(Cronbach's alpha = .844)	
McDonald's Omega	.872

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.74
Factor 2	.52
Factor 3	.15
Factor 4	-.03
Factor 5	-.07
Factor 6	-.13
Factor 7	-.14
Factor 8	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
cogself1	0.83	.004	0.82	0.83
cogself2	0.50	.008	0.48	0.51
cogself3	0.56	.007	0.54	0.57
cogself4	0.75	.005	0.74	0.76
cogself5	0.82	.004	0.81	0.83
cogself6	0.66	.006	0.64	0.67
cogself7	0.62	.007	0.61	0.63
cogself8	0.67	.006	0.66	0.68

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
cogself1	2.8	0.9	1	4	10443
cogself2	2.6	0.8	1	4	10290
cogself3	2.7	0.9	1	4	10324
cogself4	2.9	0.8	1	4	10423
cogself5	2.8	0.9	1	4	10428
cogself6	2.9	0.8	1	4	10432
cogself7	2.7	0.8	1	4	10271
cogself8	2.7	0.8	1	4	10278

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
cogself1	2.85	-4.53	-1.62	2.57
cogself2	1.13	-2.48	-0.26	2.42
cogself3	1.29	-2.66	-0.59	2.07
cogself4	2.17	-3.98	-1.53	1.87
cogself5	2.75	-4.35	-1.27	2.61
cogself6	1.67	-3.68	-1.26	1.58
cogself7	1.56	-3.22	-0.69	2.66
cogself8	1.77	-3.44	-0.88	2.53

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	943	88	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	46	14	.000
Strong Invariance (Plus Equal Intercepts)	495	14	.000
Strict Invariance (Plus Equal Error Variances)	321	14	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.998
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.999

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
cogself_fs	0.0	0.9	-2.8	2.3	10496

Share of Cases with Imputed Missing Values: 3.2%
(Equivalence of Scores from Robust MLMV: CD = .998)
(Equivalence of Scores from Two-Step Approach: CD = .983)

Scale: Cogn. Activation: Potential Solutions and Argumentation

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	332	2	.000
	Baseline vs. Saturated	19997	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.125
	90% Confidence Interval: Lower Bound	.114
	90% Confidence Interval: Upper Bound	.137
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	85451
	Bayesian Information Criterion (BIC)	85538
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.984
	Tucker-Lewis Index (TLI)	.951
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.023
	Coefficient of Determination (CD)	.878

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.864
(Cronbach's alpha = .825)	
McDonald's Omega	.865

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.37
Factor 2	-.02
Factor 3	-.08
Factor 4	-.15

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
cogself1	0.83	.004	0.83	0.84
cogself4	0.75	.005	0.74	0.76
cogself5	0.86	.004	0.85	0.86
cogself6	0.69	.006	0.68	0.71

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
cogself1	2.8	0.9	1	4	10443
cogself4	2.9	0.8	1	4	10423
cogself5	2.8	0.9	1	4	10428
cogself6	2.9	0.8	1	4	10432

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
cogself1	2.72	-4.62	-1.74	2.31
cogself4	2.19	-4.18	-1.71	1.71
cogself5	3.03	-4.91	-1.50	2.61
cogself6	1.91	-4.07	-1.49	1.53

Scale: Cogn. Activation: Potential Solutions and Argumentation (cont.) Math Module

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	351	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	24	6	.000
Strong Invariance (Plus Equal Intercepts)	110	6	.000
Strict Invariance (Plus Equal Error Variances)	105	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.997
Italian vs. German Language Version	.997

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.995

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
cogself1_fs	-0.1	0.9	-2.4	1.7	10467

Share of Cases with Imputed Missing Values: 0.9%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .985)

Scale: Cogn. Activation: Strategies and Learning From Mistakes

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	1037	2	.000
	Baseline vs. Saturated	12679	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.224
	90% Confidence Interval: Lower Bound	.212
	90% Confidence Interval: Upper Bound	.235
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	90475
	Bayesian Information Criterion (BIC)	90562
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.918
	Tucker-Lewis Index (TLI)	.755
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.061
	Coefficient of Determination (CD)	.816

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.788
(Cronbach's alpha = .743)	
McDonald's Omega	.787

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.84
Factor 2	.12
Factor 3	-.18
Factor 4	-.17

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
cogself2	0.60	.008	0.59	0.62
cogself3	0.58	.008	0.56	0.59
cogself7	0.76	.006	0.75	0.78
cogself8	0.81	.006	0.80	0.82

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
cogself2	2.6	0.8	1	4	10290
cogself3	2.7	0.9	1	4	10324
cogself7	2.7	0.8	1	4	10271
cogself8	2.7	0.8	1	4	10278

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
cogself2	1.45	-2.80	-0.35	2.62
cogself3	1.36	-2.79	-0.64	2.10
cogself7	2.13	-3.95	-0.89	3.12
cogself8	2.37	-4.28	-1.14	3.01

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	402	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	70	6	.000
Strong Invariance (Plus Equal Intercepts)	151	6	.000
Strict Invariance (Plus Equal Error Variances)	124	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.962
Italian vs. German Language Version	.975

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.936

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
cogself2_fs	0.0	0.9	-2.4	2.1	10334

Share of Cases with Imputed Missing Values: 0.9%
(Equivalence of Scores from Robust MLMV: CD = .996)
(Equivalence of Scores from Two-Step Approach: CD = .985)

Scale: Teacher: Classroom Management (Disturbances/Disorder) Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	16993	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	63509
	Bayesian Information Criterion (BIC)	63574
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.892

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.882
(Cronbach's alpha = .842)	
McDonald's Omega	.883

Test of (One-)Dimensionality (Parallel Analysis)
Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.02
Factor 2	-.09
Factor 3	-.12

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
classman1	0.79	.005	0.78	0.80
classman2	0.85	.004	0.84	0.85
classman3	0.90	.004	0.89	0.90

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
classman1	2.4	0.9	1	4	10313
classman2	2.4	0.9	1	4	10295
classman3	2.3	0.9	1	4	10272

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
classman1	2.48	-3.02	0.54	3.83
classman2	3.05	-3.19	0.28	4.06
classman3	3.96	-3.53	0.98	5.59

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	267	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	6	4	.169
Strong Invariance (Plus Equal Intercepts)	58	4	.000
Strict Invariance (Plus Equal Error Variances)	13	4	.010

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.999
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.999

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
classman_fs	0.0	0.9	-1.7	2.0	10343

Share of Cases with Imputed Missing Values: 0.9%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .992)

Scale: Teacher: Individual Learning Support

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	121	5	.000
	Baseline vs. Saturated	42736	10	.000

2)	**Root Mean Squared Error (RMSEA)**	.047
	90% Confidence Interval: Lower Bound	.040
	90% Confidence Interval: Upper Bound	.055
	Probability RMSEA <= 0.05	.730
3)	**Akaike's Information Criterion (AIC)**	94824
	Bayesian Information Criterion (BIC)	94932
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.997
	Tucker-Lewis Index (TLI)	.995
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.007
	Coefficient of Determination (CD)	.936

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.935
(Cronbach's alpha = .907)	
McDonald's Omega	.935

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.63
Factor 2	-.04
Factor 3	-.06
Factor 4	-.05
Factor 5	-.06

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
indsup1	0.86	.003	0.85	0.86
indsup2	0.89	.003	0.88	0.89
indsup3	0.87	.003	0.87	0.88
indsup4	0.87	.003	0.86	0.87
indsup5	0.82	.004	0.81	0.83

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
indsup1	2.7	0.9	1	4	10434
indsup2	3.0	0.9	1	4	10436
indsup3	2.8	0.9	1	4	10464
indsup4	2.8	0.9	1	4	10439
indsup5	2.9	0.9	1	4	10423

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
indsup1	3.14	-4.44	-1.26	2.84
indsup2	3.72	-5.69	-2.62	1.91
indsup3	3.43	-4.89	-1.86	2.64
indsup4	3.29	-4.42	-1.53	2.12
indsup5	2.74	-4.43	-1.76	2.14

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	515	40	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	35	8	.000
Strong Invariance (Plus Equal Intercepts)	196	8	.000
Strict Invariance (Plus Equal Error Variances)	57	8	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.999
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
indsup_fs	0.0	0.9	-2.2	1.6	10486

Share of Cases with Imputed Missing Values: 1.0%
(Equivalence of Scores from Robust MLMV: CD = 1.00)
(Equivalence of Scores from Two-Step Approach: CD = .981)

Scale: Situational Interest

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	11000	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			76347
	Bayesian Information Criterion (BIC)			76413
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.834

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.806
(Cronbach's alpha = .757)	
McDonald's Omega	.810

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.60
Factor 2	-.10
Factor 3	-.17

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
intsit1	0.75	.006	0.73	0.76
intsit2	0.68	.007	0.67	0.70
intsit3	0.86	.006	0.85	0.87

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
intsit1	2.6	0.9	1	4	10891
intsit2	2.3	0.9	1	4	10836
intsit3	2.4	0.9	1	4	10897

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
intsit1	2.09	-3.06	-0.39	3.19
intsit2	1.82	-1.86	0.46	3.11
intsit3	3.24	-2.76	0.54	4.35

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	801	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	282	4	.000
Strong Invariance (Plus Equal Intercepts)	61	4	.000
Strict Invariance (Plus Equal Error Variances)	251	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.974
French vs. Italian Language Version	.999
Italian vs. German Language Version	.983

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.998
Language: French	.971
Language: Italian	.995

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
intsit_fs	0.0	0.9	-1.7	2.0	10926

Share of Cases with Imputed Missing Values: 1.2%
(Equivalence of Scores from Robust MLMV: CD = .996)
(Equivalence of Scores from Two-Step Approach: CD = .988)

Scale: Perceived Autonomy Support

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	10030	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			72281
	Bayesian Information Criterion (BIC)			72346
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.809

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.799
(Cronbach's alpha = .748)	
McDonald's Omega	.800

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.55
Factor 2	-.13
Factor 3	-.17

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
persuppauto1	0.74	.006	0.73	0.76
persuppauto2	0.82	.006	0.81	0.83
persuppauto3	0.70	.007	0.69	0.72

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
persuppauto1	2.7	0.9	1	4	10665
persuppauto2	2.9	0.9	1	4	10627
persuppauto3	3.0	0.8	1	4	10655

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
persuppauto1	2.02	-3.46	-0.78	2.16
persuppauto2	2.67	-4.43	-1.76	2.02
persuppauto3	1.88	-4.13	-1.81	1.12

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	229	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	34	4	.000
Strong Invariance (Plus Equal Intercepts)	142	4	.000
Strict Invariance (Plus Equal Error Variances)	28	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.996
French vs. Italian Language Version	.994
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.993
Language: Italian	.993

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
persuppauto_fs	0.0	0.9	-2.2	1.5	10674

Share of Cases with Imputed Missing Values: 0.5%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .987)

Scale: Perceived Competence Support

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	19504	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	61112
	Bayesian Information Criterion (BIC)	61178
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.951

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.888
(Cronbach's alpha = .842)	
McDonald's Omega	.892

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.09
Factor 2	-.03
Factor 3	-.13

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
persuppcomp1	0.97	.003	0.96	0.98
persuppcomp2	0.77	.005	0.77	0.78
persuppcomp3	0.82	.004	0.81	0.83

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
persuppcomp1	2.9	0.8	1	4	10639
persuppcomp2	2.7	0.9	1	4	10639
persuppcomp3	3.0	0.8	1	4	10645

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
persuppcomp1	4.74	-7.76	-3.07	3.05
persuppcomp2	2.29	-3.63	-0.99	2.34
persuppcomp3	2.73	-5.44	-2.51	1.35

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	281	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	61	4	.000
Strong Invariance (Plus Equal Intercepts)	124	4	.000
Strict Invariance (Plus Equal Error Variances)	43	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.998
Italian vs. German Language Version	.997

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	.982

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
persuppcomp_fs	0.0	0.9	-2.2	1.5	10665

Share of Cases with Imputed Missing Values: 0.5%
(Equivalence of Scores from Robust MLMV: CD = .994)
(Equivalence of Scores from Two-Step Approach: CD = .953)

Scale: Perceived Social Relatedness

Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	15653	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	69393
	Bayesian Information Criterion (BIC)	69459
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.886

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.858
(Cronbach's alpha = .814)	
McDonald's Omega	.862

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.90
Factor 2	-.08
Factor 3	-.13

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
persocincl1	0.89	.004	0.88	0.89
persocincl2	0.70	.006	0.69	0.71
persocincl3	0.87	.004	0.86	0.88

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
persocincl1	2.7	0.9	1	4	10635
persocincl2	2.7	0.9	1	4	10640
persocincl3	2.4	0.9	1	4	10632

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
persocincl1	3.65	-4.81	-1.00	3.37
persocincl2	1.82	-3.18	-0.78	2.11
persocincl3	3.34	-2.89	0.28	4.36

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1205	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	74	4	.000
Strong Invariance (Plus Equal Intercepts)	745	4	.000
Strict Invariance (Plus Equal Error Variances)	216	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.993
French vs. Italian Language Version	.993
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.992
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
persocincl_fs	0.0	0.9	-1.9	1.8	10684

Share of Cases with Imputed Missing Values: 0.9%
(Equivalence of Scores from Robust MLMV: CD = .996)
(Equivalence of Scores from Two-Step Approach: CD = .987)

Scale: Teacher: Instruction Quality

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	9348	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			71991
	Bayesian Information Criterion (BIC)			72056
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.829

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.765
(Cronbach's alpha = .712)	
McDonald's Omega	.780

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.47
Factor 2	-.08
Factor 3	-.18

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
instqual1	0.80	.007	0.79	0.82
instqual2	0.85	.007	0.84	0.87
instqual3	0.53	.008	0.51	0.54

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
instqual1	2.8	0.9	1	4	10426
instqual2	2.8	0.8	1	4	10285
instqual3	2.6	0.9	1	4	10266

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
instqual1	2.52	-3.80	-1.28	1.99
instqual2	3.09	-4.94	-1.54	3.53
instqual3	1.15	-2.11	-0.25	2.18

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	432	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	31	4	.000
Strong Invariance (Plus Equal Intercepts)	310	4	.000
Strict Invariance (Plus Equal Error Variances)	21	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.996
French vs. Italian Language Version	.999
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	.996

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
instqual_fs	0.0	0.9	-2.0	1.7	10473

Share of Cases with Imputed Missing Values: 2.6%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .988)

Scale: ICT Interest

Math Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	15929	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			69317
	Bayesian Information Criterion (BIC)			69383
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.884

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.855
(Cronbach's alpha = .797)	
McDonald's Omega	.860

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.88
Factor 2	-.09
Factor 3	-.13

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
ictmot2	0.69	.006	0.68	0.71
ictmot3	0.88	.004	0.87	0.89
ictmot4	0.87	.004	0.86	0.88

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
ictmot2	3.2	0.7	1	4	11068
ictmot3	2.4	1.0	1	4	11065
ictmot4	2.8	0.9	1	4	11060

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
ictmot2	1.77	-4.71	-2.62	0.94
ictmot3	3.41	-3.34	0.41	3.52
ictmot4	3.42	-4.79	-1.57	2.83

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	408	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	69	4	.000
Strong Invariance (Plus Equal Intercepts)	95	4	.000
Strict Invariance (Plus Equal Error Variances)	34	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.995
French vs. Italian Language Version	.997
Italian vs. German Language Version	.995

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.994
Language: Italian	.892

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
ictintr_fs	0.0	0.9	-2.1	1.6	11071

Share of Cases with Imputed Missing Values: 0.2%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .992)

Scale: # Math Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	20861	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	68148
	Bayesian Information Criterion (BIC)	68214
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.912

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.896
(Cronbach's alpha = .849)	
McDonald's Omega	.898

Test of (One-)Dimensionality (Parallel Analysis)
Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.12
Factor 2	-.08
Factor 3	-.10

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
ictmot6	0.78	.004	0.77	0.79
ictmot7	0.90	.003	0.89	0.90
ictmot8	0.91	.003	0.90	0.91

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
ictmot6	2.9	0.9	1	4	11064
ictmot7	2.2	0.9	1	4	11057
ictmot8	2.4	0.9	1	4	11058

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
ictmot6	2.31	-4.15	-1.43	1.80
ictmot7	3.82	-2.56	1.99	5.06
ictmot8	4.06	-3.74	0.04	4.72

Scale: #NV

Math Module

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	628	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	82	4	.000
Strong Invariance (Plus Equal Intercepts)	47	4	.000
Strict Invariance (Plus Equal Error Variances)	170	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.996
French vs. Italian Language Version	.987
Italian vs. German Language Version	.997

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.996

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
ictscon_fs	0.0	0.9	-1.8	1.8	11067

Share of Cases with Imputed Missing Values: 0.2%

(Equivalence of Scores from Robust MLMV: CD = .997)

(Equivalence of Scores from Two-Step Approach: CD = .989)

Scale: Academic Self-Concept

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	31794	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	111791
	Bayesian Information Criterion (BIC)	111863
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.884

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.856
(Cronbach's alpha = .795)	
McDonald's Omega	.860

Test of (One-)Dimensionality (Parallel Analysis)
Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.89
Factor 2	-.08
Factor 3	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
scacad1	0.70	.004	0.70	0.71
scacad2	0.89	.003	0.89	0.90
scacad3	0.85	.003	0.84	0.86

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
scacad1	2.9	0.7	1	4	22202
scacad2	2.9	0.7	1	4	22175
scacad3	2.9	0.7	1	4	22168

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
scacad1	1.87	-4.54	-1.94	2.37
scacad2	3.96	-7.57	-2.86	3.92
scacad3	3.05	-6.36	-2.61	3.41

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1571	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	76	4	.000
Strong Invariance (Plus Equal Intercepts)	768	4	.000
Strict Invariance (Plus Equal Error Variances)	427	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.998
French vs. Italian Language Version	.999
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.999
Language: French	.987
Language: Italian	.996

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
scacad_fs	0.0	0.9	-2.5	1.7	22210

Share of Cases with Imputed Missing Values: 0.3%
(Equivalence of Scores from Robust MLMV: CD = .997)
(Equivalence of Scores from Two-Step Approach: CD = .986)

Scale: Verbal Self-Concept

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	32226	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	128063
	Bayesian Information Criterion (BIC)	128135
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.888

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.856
(Cronbach's alpha = .795)	
McDonald's Omega	.861

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.90
Factor 2	-.08
Factor 3	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
scverb1 *	0.70	0.00	0.69	0.70
scverb2	0.90	0.00	0.89	0.90
scverb3	0.86	0.00	0.85	0.86

*** Note:** Reversed Categories for Item scverb1

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
scverb1	3.2	0.8	1	4	22196
scverb2	2.8	0.8	1	4	22173
scverb3	2.9	0.8	1	4	22171

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
scverb1 *	1.84	-4.49	-2.24	0.34
scverb2	3.52	-6.01	-1.79	3.39
scverb3	2.89	-5.94	-2.37	2.79

*** Note:** Reversed Categories for Item Scverb1

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	621	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	30	4	.000
Strong Invariance (Plus Equal Intercepts)	58	4	.000
Strict Invariance (Plus Equal Error Variances)	215	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.989
Italian vs. German Language Version	.986

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	.998

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
scverb_fs	0.0	0.9	-2.4	1.6	22205

Share of Cases with Imputed Missing Values: 0.3%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .988)

Scale: Global Self-Esteem

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	13531	20	.000
	Baseline vs. Saturated	44683	28	.000

2)	**Root Mean Squared Error (RMSEA)**	.247
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	229076
	Bayesian Information Criterion (BIC)	229252
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.697
	Tukey-Lewis Index (TLI)	.576
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.143
	Coefficient of Determination (CD)	.889

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.861
(Cronbach's alpha = .822)	
McDonald's Omega	.854

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.60
Factor 2	1.09
Factor 3	.07
Factor 4	-.05
Factor 5	-.09
Factor 6	-.10
Factor 7	-.12
Factor 8	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
sele1	0.64	.007	0.63	0.66
sele2	0.52	.008	0.50	0.53
sele3	0.45	.009	0.44	0.47
sele4	0.50	.008	0.48	0.52
seld1 *	0.85	.004	-0.86	-0.84
seld3 *	0.74	.005	-0.75	-0.73
seld4 *	0.65	.006	-0.67	-0.64
seld5 *	0.80	.004	-0.81	-0.79

*** Note:** Reversed Categories for "Seld"-Items

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
sele1	3.9	0.9	1	5	11053
sele2	4.1	0.8	1	5	11032
sele3	3.9	0.9	1	5	11032
sele4	3.7	1.1	1	5	11017
seld1	2.2	1.2	1	5	11038
seld3	2.7	1.2	1	5	11023
seld4	2.8	1.3	1	5	10997
seld5	1.9	1.2	1	5	11010

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	3840	88	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	93	14	.000
Strong Invariance (Plus Equal Intercepts)	2088	14	.000
Strict Invariance (Plus Equal Error Variances)	321	14	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.997
French vs. Italian Language Version	.992
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.989
Language: Italian	.997

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
sel_fs	0.0	0.5	-1.8	0.8	11063

Share of Cases with Imputed Missing Values: 1.2%

(Equivalence of Scores from Robust MLMV: CD = .996)

Scale: Positive Global Self-Esteem

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	232	2	.000
	Baseline vs. Saturated	18398	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.102
	90% Confidence Interval: Lower Bound	.091
	90% Confidence Interval: Upper Bound	.113
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	98727
	Bayesian Information Criterion (BIC)	98815
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.987
	Tucker-Lewis Index (TLI)	.962
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.019
	Coefficient of Determination (CD)	.857

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.848
(Cronbach's alpha = .801)	
McDonald's Omega	.849

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.21
Factor 2	-.06
Factor 3	-.08
Factor 4	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
sele1	0.72	.006	0.71	0.73
sele2	0.83	.004	0.82	0.84
sele3	0.79	.005	0.78	0.80
sele4	0.71	.006	0.70	0.72

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
sele1	3.9	0.9	1	5	11053
sele2	4.1	0.8	1	5	11032
sele3	3.9	0.9	1	5	11032
sele4	3.7	1.1	1	5	11017

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1216	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	12	6	.062
Strong Invariance (Plus Equal Intercepts)	785	6	.000
Strict Invariance (Plus Equal Error Variances)	128	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	1.000
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
sele_fs	0.0	0.6	-2.5	0.9	11057

Share of Cases with Imputed Missing Values: 0.6%
(Equivalence of Scores from Robust MLMV: CD = .996)

Scale: Negative Global Self-Esteem

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	394	2	.000
	Baseline vs. Saturated	21952	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.133
	90% Confidence Interval: Lower Bound	.122
	90% Confidence Interval: Upper Bound	.144
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	121631
	Bayesian Information Criterion (BIC)	121719
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.982
	Tucker-Lewis Index (TLI)	.946
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.026
	Coefficient of Determination (CD)	.890

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.866
(Cronbach's alpha = .822)	
McDonald's Omega	.868

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.38
Factor 2	.01
Factor 3	-.13
Factor 4	-.10

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
seld1	0.89	.003	0.89	0.90
seld3	0.78	.004	0.77	0.79
seld4	0.67	.006	0.66	0.68
seld5	0.80	.004	0.79	0.81

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
seld1	2.2	1.2	1	5	11038
seld3	2.7	1.2	1	5	11023
seld4	2.8	1.3	1	5	10997
seld5	1.9	1.2	1	5	11010

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	2592	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	40	6	.000
Strong Invariance (Plus Equal Intercepts)	1565	6	.000
Strict Invariance (Plus Equal Error Variances)	193	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	1.000
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.991
Language: Italian	.989

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
seld_fs	0.0	1.0	-1.3	2.6	11057

Share of Cases with Imputed Missing Values: 0.9%
(Equivalence of Scores from Robust MLMV: CD = .996)

Scale: General Perceived Self-Efficacy Scale (GSES) Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	49	2	.000
	Baseline vs. Saturated	16703	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.046
	90% Confidence Interval: Lower Bound	.035
	90% Confidence Interval: Upper Bound	.058
	Probability RMSEA <= 0.05	.690
3)	**Akaike's Information Criterion (AIC)**	73468
	Bayesian Information Criterion (BIC)	73556
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.997
	Tucker-Lewis Index (TLI)	.992
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.009
	Coefficient of Determination (CD)	.840

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.838
(Cronbach's alpha = .776)	
McDonald's Omega	.839

Test of (One-)Dimensionality (Parallel Analysis)
Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.13
Factor 2	-.09
Factor 3	-.11
Factor 4	-.12

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
seef1	0.73	.006	0.72	0.74
seef2	0.78	.005	0.77	0.79
seef3	0.76	.005	0.75	0.77
seef4	0.73	.006	0.72	0.74

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
seef1	3.1	0.6	1	4	11019
seef2	3.1	0.7	1	4	11014
seef3	2.8	0.7	1	4	11010
seef4	3.0	0.7	1	4	11014

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
seef1	2.04	-5.94	-3.08	2.22
seef2	2.36	-6.19	-2.91	1.92
seef3	2.17	-5.12	-1.38	2.66
seef4	2.05	-5.53	-1.96	2.28

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	715	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	26	6	.000
Strong Invariance (Plus Equal Intercepts)	290	6	.000
Strict Invariance (Plus Equal Error Variances)	179	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.997
Italian vs. German Language Version	.997

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.997
Language: Italian	.994

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
seef_fs	0.0	0.9	-3.0	1.8	11026

Share of Cases with Imputed Missing Values: 0.2%
(Equivalence of Scores from Robust MLMV: CD = .996)
(Equivalence of Scores from Two-Step Approach: CD = .988)

Scale: Perseverance

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	18182	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	168695
	Bayesian Information Criterion (BIC)	168767
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.825

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.767
(Cronbach's alpha = .731)	
McDonald's Omega	.775

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.43
Factor 2	-.09
Factor 3	-.20

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
persev1	0.67	.005	0.66	0.68
persev2	0.87	.005	0.86	0.88
persev3	0.64	.005	0.63	0.65

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
persev1	3.5	0.9	1	5	22268
persev2	3.4	1.0	1	5	22269
persev3	2.9	1.0	1	5	22265

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	2678	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	79	4	.000
Strong Invariance (Plus Equal Intercepts)	1498	4	.000
Strict Invariance (Plus Equal Error Variances)	207	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.997
French vs. Italian Language Version	.999
Italian vs. German Language Version	.994

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.998
Language: French	.990
Language: Italian	.989

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
persev_fs	0.0	0.5	-1.5	1.1	22280

Share of Cases with Imputed Missing Values: 0.1%
(Equivalence of Scores from Robust MLMV: CD = .997)

Composit Descriptives

Big Five Inventory	Likelihood-Ratio Tests Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
Big Five: Extraversion	big5_e_comp	3.3	0.9	1	5	11003
Big Five: Agreeableness	big5_a_comp	3.4	0.7	1	5	11003
Big Five: Conscientiousness	big5_c_comp	3.2	0.8	1	5	11003
Big Five: Neuroticism	big5_n_comp	2.9	0.9	1	5	11003
Big Five: Openness	big5_o_comp	3.3	0.9	1	5	11003
Share of Cases with Imputed Missing Values:		1.3%				

Item Descriptives

Big Five Inventory	Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.	
Big Five: Extraversion	bigfive1	3.1	1.1	1	5	10983	*
	bigfive6	3.5	1.0	1	5	10953	
Big Five: Agreeableness	bigfive2	3.2	1.1	1	5	10976	
	bigfive7	3.3	1.0	1	5	10951	*
	bigfive11	3.8	1.1	1	5	10948	
Big Five: Conscientiousness	bigfive3	2.8	1.1	1	5	10968	*
	bigfive8	3.5	0.9	1	5	10956	
Big Five: Neuroticism	bigfive4	2.8	1.1	1	5	10970	*
	bigfive9	2.9	1.1	1	5	10963	
Big Five: Openness	bigfive5	3.0	1.4	1	5	10966	*
	bigfive10	3.7	1.1	1	5	10959	

* Items reversed for composit calculation (see Rammstedt et al., 2007)

Scale: Extrinsic Value Orientation

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	4765	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	67121
	Bayesian Information Criterion (BIC)	67187
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.673

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.662
(Cronbach's alpha = .564)	
McDonald's Omega	.664

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.01
Factor 2	-.14
Factor 3	-.19

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
vawe1	0.70	.010	0.68	0.72
vawe2	0.62	.010	0.60	0.64
vawe4	0.57	.010	0.55	0.59

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
vawe1	3.2	0.7	1	4	11086
vawe2	3.6	0.6	1	4	11079
vawe4	2.9	0.9	1	4	11075

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
vawe1	1.82	-5.53	-2.59	0.97
vawe2	1.43	-5.46	-4.02	-1.02
vawe4	1.21	-3.31	-1.01	1.37

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	206	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	24	4	.000
Strong Invariance (Plus Equal Intercepts)	51	4	.000
Strict Invariance (Plus Equal Error Variances)	55	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.993
French vs. Italian Language Version	.980
Italian vs. German Language Version	.995

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.999
Language: French	.986
Language: Italian	.951

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
vawe_fs	0.0	0.8	-2.8	1.2	11091

Share of Cases with Imputed Missing Values: 0.2%
(Equivalence of Scores from Robust MLMV: CD = .996)
(Equivalence of Scores from Two-Step Approach: CD = .975)

Scale: Intrinsic Value Orientation

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	10457	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	57140
	Bayesian Information Criterion (BIC)	57206
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.824

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.796
(Cronbach's alpha = .713)	
McDonald's Omega	.800

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.53
Factor 2	-.10
Factor 3	-.17

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
vawi1	0.73	.006	0.71	0.74
vawi2	0.85	.006	0.84	0.87
vawi5	0.68	.007	0.67	0.69

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
vawi1	3.2	0.7	1	4	11087
vawi2	3.5	0.6	1	4	11082
vawi5	3.5	0.7	1	4	11077

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
vawi1	1.87	-5.19	-2.64	1.04
vawi2	3.31	-8.90	-6.25	-0.61
vawi5	1.68	-5.48	-3.72	-0.36

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	336	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	6	4	.165
Strong Invariance (Plus Equal Intercepts)	125	4	.000
Strict Invariance (Plus Equal Error Variances)	124	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.999
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
vawi_fs	0.0	0.8	-3.0	1.1	11092

Share of Cases with Imputed Missing Values: 0.2%
(Equivalence of Scores from Robust MLMV: CD = .991)
(Equivalence of Scores from Two-Step Approach: CD = .962)

Composit: Family Value Orientation

Composit Descriptives

Family Value Orientation	Likelihood-Ratio Tests Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
Family Value Orientation	vafa_comp	3.1	0.8	1	4	11083
Share of Cases with Imputed Missing Values:		0.2%				

Item Descriptives

Family Value Orientation	Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
Family Value Orientation	vafa1	3.3	0.8	1	4	11077
	vafa2	3.0	0.9	1	4	11071

Scale: Interest in Reading

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	44643	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	153979
	Bayesian Information Criterion (BIC)	154051
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.924

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.906
(Cronbach's alpha = .864)	
McDonald's Omega	.907

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.19
Factor 2	-.07
Factor 3	-.11

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
intrea1	0.86	.002	0.85	0.86
intrea2	0.94	.002	0.93	0.94
intrea3	0.83	.003	0.82	0.83

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
intrea1	2.2	1.0	1	4	22180
intrea2	2.1	1.1	1	4	22178
intrea3	2.3	1.1	1	4	22165

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
intrea1	3.03	-1.81	0.96	3.55
intrea2	5.35	-1.65	2.08	5.65
intrea3	2.63	-1.67	0.17	2.61

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	732	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	94	4	.000
Strong Invariance (Plus Equal Intercepts)	560	4	.000
Strict Invariance (Plus Equal Error Variances)	7	4	.155

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	1.000
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.999
Language: French	.998
Language: Italian	.998

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
intrea_fs	0.0	0.9	-1.3	1.7	22200

Share of Cases with Imputed Missing Values: 0.3%
(Equivalence of Scores from Robust MLMV: CD = .997)
(Equivalence of Scores from Two-Step Approach: CD = .973)

Scale: Self-Reported Capabilities

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	1254	5	.000
	Baseline vs. Saturated	26986	10	.000

2)	**Root Mean Squared Error (RMSEA)**	.151
	90% Confidence Interval: Lower Bound	.144
	90% Confidence Interval: Upper Bound	.159
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	154479
	Bayesian Information Criterion (BIC)	154588
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.954
	Tucker-Lewis Index (TLI)	.907
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.039
	Coefficient of Determination (CD)	.880

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.877
(Cronbach's alpha = .854)	
McDonald's Omega	.877

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.85
Factor 2	.10
Factor 3	-.06
Factor 4	-.12
Factor 5	-.12

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
cap1	0.77	.005	0.76	0.78
cap2	0.79	.005	0.78	0.80
cap3	0.80	.004	0.79	0.81
cap4	0.71	.006	0.70	0.72
cap5	0.77	.005	0.76	0.78

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
cap1	5.8	1.3	1	7	10876
cap2	5.6	1.3	1	7	10863
cap3	5.8	1.3	1	7	10865
cap4	5.3	1.3	1	7	10853
cap5	5.6	1.2	1	7	10867

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	793	40	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	71	8	.000
Strong Invariance (Plus Equal Intercepts)	378	8	.000
Strict Invariance (Plus Equal Error Variances)	109	8	.000

Configural Factor Similarity	
Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.996
French vs. Italian Language Version	.998
Italian vs. German Language Version	.997

Factor Score Equivalence: Group Specific vs. Invariant Models	
Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	.999

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
cap_fs	0.0	0.9	-4.3	1.2	10900

Share of Cases with Imputed Missing Values: 0.6%
(Equivalence of Scores from Robust MLMV: CD = .998)

Scale: Instrumental Motivation

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	28969	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	144091
	Bayesian Information Criterion (BIC)	144163
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.865

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.848
(Cronbach's alpha = .796)	
McDonald's Omega	.850

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.81
Factor 2	-.10
Factor 3	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
insmot1	0.75	0.00	0.74	0.76
insmot2	0.79	0.00	0.78	0.80
insmot3	0.88	0.00	0.88	0.89

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
insmot1	2.8	0.9	1	4	22246
insmot2	2.9	0.9	1	4	22220
insmot3	3.1	0.9	1	4	22220

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
insmot1	2.05	-3.82	-0.83	2.13
insmot2	2.35	-3.90	-1.28	1.70
insmot3	3.48	-6.32	-3.28	0.89

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	347	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	29	4	.000
Strong Invariance (Plus Equal Intercepts)	136	4	.000
Strict Invariance (Plus Equal Error Variances)	55	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.997
Italian vs. German Language Version	.994

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.982

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
insmot_fs	0.0	0.9	-2.2	1.4	22265

Share of Cases with Imputed Missing Values: 0.4%
(Equivalence of Scores from Robust MLMV: CD = .996)
(Equivalence of Scores from Two-Step Approach: CD = .978)

Composit: Locus of Control

Composit Descriptives

Locus of Control	Likelihood-Ratio Tests Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
Internal Locus of Control	loc_i_comp	4.0	0.8	1	5	10933
External Locus of Control	loc_e_comp	2.5	0.9	1	5	10933
Share of Cases with Imputed Missing Values:		0.4%				

Item Descriptives

Locus of Control	Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
Internal Locus of Control	loci1	3.8	0.9	1	5	10920
	loci2	4.2	0.8	1	5	10916
External Locus of Control	loce1	2.3	1.1	1	5	10905
	loce2	2.7	1.1	1	5	10904

Scale: Intrinsic Achievement Motivation

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	12995	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	152039
	Bayesian Information Criterion (BIC)	152111
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.795

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.703
(Cronbach's alpha = .652)	
McDonald's Omega	.718

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.19
Factor 2	-.08
Factor 3	-.22

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
achmot2	0.54	.006	0.52	0.55
achmot4	0.62	.006	0.60	0.63
achmot6	0.86	.007	0.85	0.87

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
achmot2	3.0	0.8	1	4	22249
achmot4	2.8	0.8	1	4	22242
achmot6	2.6	0.9	1	4	22239

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
achmot2	1.16	-3.58	-1.45	1.12
achmot4	1.47	-3.30	-0.89	2.11
achmot6	2.88	-4.12	-0.77	3.70

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1286	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	14	4	.007
Strong Invariance (Plus Equal Intercepts)	956	4	.000
Strict Invariance (Plus Equal Error Variances)	141	4	.000

Configural Factor Similarity	
Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.993
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models	
Coefficient of Determination	CD
Language: German	.999
Language: French	.999
Language: Italian	.990

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
achmoti_fs	0.0	0.9	-2.2	1.8	22262

Share of Cases with Imputed Missing Values: 0.2%
(Equivalence of Scores from Robust MLMV: CD = .994)
(Equivalence of Scores from Two-Step Approach: CD = .982)

Scale: Extrinsic Achievement Motivation

Core Questionnaire

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	12774	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	148710
	Bayesian Information Criterion (BIC)	148782
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.792

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.648
(Cronbach's alpha = .589)	
McDonald's Omega	.690

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.14
Factor 2	-.04
Factor 3	-.22

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
achmot1	0.33	.007	0.32	0.34
achmot3	0.73	.009	0.72	0.75
achmot5	0.85	.009	0.83	0.86

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
achmot1	3.2	0.7	1	4	22263
achmot3	1.8	0.8	1	4	22239
achmot5	1.9	0.9	1	4	22235

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
achmot1	0.58	-3.66	-2.13	0.51
achmot3	2.18	-0.50	2.38	5.22
achmot5	2.49	-0.62	2.16	5.11

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1767	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	36	4	.000
Strong Invariance (Plus Equal Intercepts)	954	4	.000
Strict Invariance (Plus Equal Error Variances)	211	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.982
French vs. Italian Language Version	.995
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.979
Language: French	.961
Language: Italian	.993

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
achmote_fs	0.0	0.8	-1.3	2.3	22266

Share of Cases with Imputed Missing Values: 0.2%
(Equivalence of Scores from Robust MLMV: CD = .990)
(Equivalence of Scores from Two-Step Approach: CD = .981)

Composit: Learning Effort

Composit Descriptives

Learning Effort	Likelihood-Ratio Tests Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
Learning Effort	effper_comp	2.8	0.8	1	4	22265
Share of Cases with Imputed Missing Values:		0.2%				

Item Descriptives

Learning Effort	Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
Learning Effort	effper1	2.7	0.8	1	4	22243
	effper4	2.9	0.9	1	4	22249

Scale: Absenteeism (Truancy) Core Questionnaire

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	30122	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			84033
	Bayesian Information Criterion (BIC)			84105
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.923

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.819
(Cronbach's alpha = .648)	
McDonald's Omega	.837

Test of (One-)Dimensionality (Parallel Analysis)
Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.77
Factor 2	-.03
Factor 3	-.14

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
truancy1	0.84	.004	0.83	0.85
truancy2	0.95	.004	0.94	0.96
truancy3	0.56	.005	0.55	0.57

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
truancy1	1.1	0.4	1	4	22242
truancy2	1.2	0.5	1	4	22245
truancy3	1.5	0.8	1	4	22251

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
truancy1	3.27	4.85	7.51	8.62
truancy2	4.63	5.31	8.79	10.99
truancy3	1.16	0.54	2.44	3.49

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	2001	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	38	4	.000
Strong Invariance (Plus Equal Intercepts)	734	4	.000
Strict Invariance (Plus Equal Error Variances)	680	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.998
Italian vs. German Language Version	1.000

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.997
Language: French	.988
Language: Italian	.954

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
truancy_fs	0.0	0.7	-0.5	2.8	22254

Share of Cases with Imputed Missing Values: 0.1%
(Equivalence of Scores from Robust MLMV: CD = .995)
(Equivalence of Scores from Two-Step Approach: CD = .780)

Composit: Reading Interest

Core Questionnaire

Composit Descriptives

Parents: Reading Interest	Likelihood-Ratio Tests Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
Parents: Reading Interest	joyreadp_comp	3.1	0.8	1	4	10531
Share of Cases with Imputed Missing Values: (Including "Don't Know"-Answers for One Parent)		8.5%				

Item Descriptives

Parents: Reading Interest	Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
Parents: Reading Interest	joyreadm	3.4	0.9	1	4	10370
	joyreadf	2.8	1.1	1	4	9801

Composit: Emotional Closeness to Parents

Monitoring Module

Composit Descriptives

Emotional Closeness to Parents	Likelihood-Ratio Tests Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
Emotional Closeness to Parents	closep_comp	4.2	0.8	1	5	10824
Share of Cases with Imputed Missing Values:		3.1%				

Item Descriptives

Emotional Closeness to Parents	Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
Emotional Closeness to Parents	closef	4.1	1.1	1	5	10552
	closem	4.4	0.9	1	5	10763

Scale: Parental Pressure to Achieve

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	271	2	.000
	Baseline vs. Saturated	14102	6	.000

2)	**Root Mean Squared Error (RMSEA)**	.112
	90% Confidence Interval: Lower Bound	.101
	90% Confidence Interval: Upper Bound	.124
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	97596
	Bayesian Information Criterion (BIC)	97683
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.981
	Tucker-Lewis Index (TLI)	.943
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.024
	Coefficient of Determination (CD)	.822

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.814
(Cronbach's alpha = .756)	
McDonald's Omega	.815

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.96
Factor 2	-.04
Factor 3	-.09
Factor 4	-.17

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
press1	0.68	.007	0.67	0.69
press2	0.70	.007	0.68	0.71
press3	0.79	.006	0.78	0.81
press4	0.72	.006	0.71	0.74

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
press1	2.2	1.0	1	4	10638
press2	3.0	0.9	1	4	10634
press3	3.1	0.8	1	4	10631
press4	2.8	0.9	1	4	10641

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
press1	1.61	-1.41	0.66	2.95
press2	1.79	-3.60	-1.78	0.78
press3	2.44	-5.21	-2.32	1.40
press4	1.94	-3.50	-1.23	1.55

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1159	28	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	26	6	.000
Strong Invariance (Plus Equal Intercepts)	590	6	.000
Strict Invariance (Plus Equal Error Variances)	48	6	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	.991
Italian vs. German Language Version	.986

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.959

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
press_fs	0.0	0.9	-2.4	1.7	10668

Share of Cases with Imputed Missing Values: 0.5%
(Equivalence of Scores from Robust MLMV: CD = .997)
(Equivalence of Scores from Two-Step Approach: CD = .984)

Scale: Family Educational Support

Monitoring Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	11883	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			101555
	Bayesian Information Criterion (BIC)			101620
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.866

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.790
(Cronbach's alpha = .751)	
McDonald's Omega	.808

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.62
Factor 2	-.06
Factor 3	-.16

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
famedsup1	0.87	.006	0.86	0.88
famedsup2	0.86	.006	0.85	0.87
famedsup3	0.53	.008	0.51	0.54

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
famedsup1	2.8	1.4	1	5	10654
famedsup2	2.6	1.4	1	5	10452
famedsup3	2.3	1.4	1	5	9257

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	268	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	26	4	.000
Strong Invariance (Plus Equal Intercepts)	216	4	.000
Strict Invariance (Plus Equal Error Variances)	17	4	.002

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.997
French vs. Italian Language Version	.993
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.997
Language: Italian	.976

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
famedsup_fs	0.0	1.1	-1.5	2.2	10732

Share of Cases with Imputed Missing Values: 15.9%
(Equivalence of Scores from Robust MLMV: CD = .998)

Scale: Cultural Communication

Monitoring Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	11409	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			95435
	Bayesian Information Criterion (BIC)			95501
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.831

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.816
(Cronbach's alpha = .767)	
McDonald's Omega	.818

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.64
Factor 2	-.11
Factor 3	-.16

Standardized Factor Loadings

Indicators *	Coef.	(SE)	[95% Conf.	Interval]
cultcom1	0.80	.006	0.79	0.81
cultcom2	0.69	.006	0.68	0.70
cultcom4	0.83	.006	0.82	0.84

*** Note:** Item Cultcom3 Excluded to Improve Scale Quality

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
cultcom1	2.9	1.4	1	5	10732
cultcom2	3.1	1.3	1	5	10723
cultcom4	3.8	1.2	1	5	10723

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	221	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	15	4	.005
Strong Invariance (Plus Equal Intercepts)	124	4	.000
Strict Invariance (Plus Equal Error Variances)	32	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.998
Italian vs. German Language Version	.996

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.997

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
cultcom_m_fs	0.0	1.0	-2.4	1.5	10745

Share of Cases with Imputed Missing Values: 0.4%

(Equivalence of Scores from Robust MLMV: CD = .997)

Scale: Social Communication

Monitoring Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	18800	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			83020
	Bayesian Information Criterion (BIC)			83086
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.893

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.893
(Cronbach's alpha = .855)	
McDonald's Omega	.893

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.08
Factor 2	-.11
Factor 3	-.10

Standardized Factor Loadings

Indicators *	Coef.	(SE)	[95% Conf.	Interval]
soccom3	0.84	.004	0.84	0.85
soccom4	0.86	.004	0.86	0.87
soccom5	0.86	.004	0.86	0.87

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
soccom3	3.9	1.2	1	5	10705
soccom4	3.8	1.2	1	5	10705
soccom5	4.0	1.2	1	5	10710

*** Note:** Items Soccom1 and Soccom2 Excluded to Improve Scale Quality

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	628	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	26	4	.000
Strong Invariance (Plus Equal Intercepts)	70	4	.000
Strict Invariance (Plus Equal Error Variances)	164	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	1.000
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.997

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
soccom_m_fs	0.0	0.9	-2.6	1.0	10728

Share of Cases with Imputed Missing Values: 0.4%

(Equivalence of Scores from Robust MLMV: CD = .997)

Scale: Perceived Social Network Support

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	1440	5	.000
	Baseline vs. Saturated	40854	10	.000

2)	**Root Mean Squared Error (RMSEA)**	.167
	90% Confidence Interval: Lower Bound	.160
	90% Confidence Interval: Upper Bound	.175
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	161342
	Bayesian Information Criterion (BIC)	161451
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.965
	Tucker-Lewis Index (TLI)	.930
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.035
	Coefficient of Determination (CD)	.943

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.922
(Cronbach's alpha = .899)	
McDonald's Omega	.923

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.48
Factor 2	.09
Factor 3	-.01
Factor 4	-.06
Factor 5	-.11

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
closupp1	0.81	.004	0.80	0.82
closupp2	0.93	.002	0.93	0.94
closupp3	0.89	.002	0.89	0.89
closupp4	0.68	.006	0.67	0.69
closupp5	0.86	.003	0.86	0.87

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
closupp1	5.2	1.7	1	7	10038
closupp2	5.5	1.7	1	7	10081
closupp3	5.6	1.6	1	7	10071
closupp4	5.0	1.7	1	7	9742
closupp5	5.4	1.8	1	7	9879

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	438	40	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	17	8	.027
Strong Invariance (Plus Equal Intercepts)	136	8	.000
Strict Invariance (Plus Equal Error Variances)	205	8	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.999
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	1.000

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
closupp_fs	0.0	1.3	-3.9	1.3	10234

Share of Cases with Imputed Missing Values: 8.3%

(Equivalence of Scores from Robust MLMV: CD = .999)

Scale: Family Wealth

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	7993	27	.000
	Baseline vs. Saturated	33591	36	.000

2)	**Root Mean Squared Error (RMSEA)**	.162
	90% Confidence Interval: Lower Bound	.159
	90% Confidence Interval: Upper Bound	.165
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	91016
	Bayesian Information Criterion (BIC)	91214
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.763
	Tucker-Lewis Index (TLI)	.683
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.077
	Coefficient of Determination (CD)	.854

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.823
(Cronbach's alpha = .612)	
McDonald's Omega	.825

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.21
Factor 2	.55
Factor 3	.28
Factor 4	.09
Factor 5	.03
Factor 6	-.01
Factor 7	-.08
Factor 8	-.19
Factor 9	-.22

Standardized Factor Loadings

Indicators *	Coef.	(SE)	[95% Conf.	Interval]
wealth1	0.74	.005	0.73	0.75
wealth2	0.60	.007	0.59	0.62
wealth4	0.78	.005	0.77	0.79
wealth5	0.62	.007	0.60	0.63
wealthn1	0.56	.008	0.54	0.57
wealthn2	0.33	.009	0.31	0.35
wealthn3	0.51	.008	0.49	0.52
wealthn4	0.47	.008	0.45	0.49
wealthn5	0.63	.007	0.62	0.65

*** Note:** Item Wealth3 Excluded to Improve Scale Quality

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
wealth1	0.9	0.3	0	1	11172
wealth2	0.9	0.3	0	1	11176
wealth4	1.0	0.1	0	1	11177
wealth5	0.6	0.5	0	1	11159
wealthn1	3.9	0.4	1	4	11171
wealthn2	2.8	0.8	1	4	11172
wealthn3	3.3	0.8	1	4	11167
wealthn4	2.7	0.8	1	4	11166
wealthn5	2.9	0.7	1	4	11171

Parameters of Generalized Structural Equation Model ([Ordinal] Logit Link]

Indicators	Coef.	Cut1	Cut2	Cut3
wealth1	1.87	-3.69		
wealth2	1.30	-2.90		
wealth4	2.00	-6.38		
wealth5	1.49	-0.72		
wealthn1	1.24	-5.96	-4.20	-3.10
wealthn2	0.63	-3.46	-0.53	1.28
wealthn3	0.93	-4.78	-1.64	-0.04
wealthn4	1.16	-3.14	-0.24	2.18
wealthn5	1.68	-6.43	-1.34	2.13

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices *	chi2	df	p > chi2
	1879	54	.000

Tests of Measurement Invariance *	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	25	8	.002
Strong Invariance (Plus Equal Intercepts)	231	8	.000
Strict Invariance (Plus Equal Error Variances)	469	8	.000

Configural Factor Similarity *	
Tucker's Congruence Coefficient	TCC
German vs. French or Italian Version	1.000

Factor Score Equivalence: Group-Specific vs. Invariant Model *	
Coefficient of Determination	CD
Language: German	1.000
Language: French/ Italian	.999

* Note: Due to sparse tables for the italian version of the scale, equivalence tests failed to converge and were reestimated with collapsed italian and french versions.

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
wealth_fs	0.0	0.8	-3.9	1.8	11185

Share of Cases with Imputed Missing Values: 0.5%
(Equivalence of Scores from Robust MLMV: CD = .816)
(Equivalence of Scores from Two-Step Approach: CD = .718)

Scale: Home Possessions: 'Classical Culture'

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	7669	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	36799
	Bayesian Information Criterion (BIC)	36864
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.808

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.708
(Cronbach's alpha = .544)	
McDonald's Omega	.732

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.26
Factor 2	-.05
Factor 3	-.20

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
cultposs1	0.74	.009	0.72	0.76
cultposs2	0.86	.010	0.84	0.87
cultposs3	0.44	.009	0.42	0.46

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
cultposs1	0.4	0.5	0	1	11131
cultposs2	0.4	0.5	0	1	11141
cultposs3	0.7	0.4	0	1	11152

Parameters of Generalized Structural Equation Model (Logit Link)

Indicators	Coef.	Cut1
cultposs1	1.89	0.75
cultposs2	3.16	0.61
cultposs3	0.86	-1.20

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices *	chi2	df	p > chi2
	3036	9	.000

Tests of Measurement Invariance *	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	66	2	.000
Strong Invariance (Plus Equal Intercepts)	583	2	.000
Strict Invariance (Plus Equal Error Variances)	229	2	.000

Configural Factor Similarity *	
Tucker's Congruence Coefficient	TCC
German vs. French or Italian Version	.984

Factor Score Equivalence: Group-Specific vs. Invariant Model *	
Coefficient of Determination	CD
Language: German	.984
Language: French/ Italian	.761

* Note: Due to sparse tables for the italian version of the scale, equivalence tests failed to converge and were reestimated with collapsed italian and french versions.

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
cultposs_fs	0.0	0.8	-1.0	1.1	11165

Share of Cases with Imputed Missing Values: 0.4%
(Equivalence of Scores from Robust MLMV: CD = .968)
(Equivalence of Scores from Two-Step Approach: CD = .958)

Scale: Cultural Activities

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	6024	14	.000
	Baseline vs. Saturated	21138	21	.000

2)	**Root Mean Squared Error (RMSEA)**	.198
	90% Confidence Interval: Lower Bound	.194
	90% Confidence Interval: Upper Bound	.202
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	179660
	Bayesian Information Criterion (BIC)	179814
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.715
	Tucker-Lewis Index (TLI)	.573
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.123
	Coefficient of Determination (CD)	.824

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.753
(Cronbach's alpha = .679)	
McDonald's Omega	.736

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	2.22
Factor 2	.80
Factor 3	.06
Factor 4	-.03
Factor 5	-.13
Factor 6	-.18
Factor 7	-.19

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
cult1	0.37	.009	0.35	0.39
cult2	0.71	.006	0.70	0.73
cult3	0.49	.009	0.48	0.51
cult4	0.79	.005	0.78	0.80
cult5	0.76	.006	0.75	0.77
cult7	0.29	.010	0.27	0.31
cult9	0.23	.010	0.21	0.25

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
cult1	2.8	1.0	1	4	10926
cult2	1.7	0.9	1	4	10920
cult3	1.6	0.8	1	4	10913
cult4	1.3	0.6	1	4	10912
cult5	1.6	0.7	1	4	10905
cult7	2.6	1.0	1	4	10910
cult9	2.5	1.2	1	4	10907

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
cult1	0.88	-2.78	-0.53	0.89
cult2	1.64	-0.22	2.22	3.73
cult3	1.16	0.41	2.46	3.62
cult4	2.06	2.26	4.32	5.57
cult5	1.87	0.18	3.22	4.87
cult7	0.69	-1.82	0.17	1.40
cult9	0.57	-0.94	0.27	1.14

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	1057	70	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	60	12	.000
Strong Invariance (Plus Equal Intercepts)	852	12	.000
Strict Invariance (Plus Equal Error Variances)	98	12	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.994
French vs. Italian Language Version	.992
Italian vs. German Language Version	.985

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	.997

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
cult_fs	0.0	0.8	-1.8	3.0	10934

Share of Cases with Imputed Missing Values: 0.6%
(Equivalence of Scores from Robust MLMV: CD = .980)
(Equivalence of Scores from Two-Step Approach: CD = .883)

Scale: Popular Cultural Activities

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	5396	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	86373
	Bayesian Information Criterion (BIC)	86439
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.737

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.679
(Cronbach's alpha = .609)	
McDonald's Omega	.690

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.08
Factor 2	-.10
Factor 3	-.21

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
cult3	0.56	.010	0.54	0.58
cult7	0.58	.010	0.56	0.60
cult9	0.81	.010	0.78	0.83

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
cult3	1.6	0.8	1	4	10913
cult7	2.6	1.0	1	4	10910
cult9	2.5	1.2	1	4	10907

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
cult3	1.16	0.41	2.48	3.60
cult7	1.28	-2.15	0.17	1.63
cult9	2.30	-1.58	0.44	1.88

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	718	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	38	4	.000
Strong Invariance (Plus Equal Intercepts)	513	4	.000
Strict Invariance (Plus Equal Error Variances)	103	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.988
French vs. Italian Language Version	.987
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	.994
Language: French	.977
Language: Italian	.954

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
cultlow_fs	0.0	0.8	-1.4	1.8	10926

Share of Cases with Imputed Missing Values: 0.3%
(Equivalence of Scores from Robust MLMV: CD = .989)
(Equivalence of Scores from Two-Step Approach: CD = .975)

Scale: Highbrow Cultural Activities

Monitoring Module

Model and Fit Statistics

1)	**Likelihood-Ratio Tests**	chi2	df	p > chi2
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	11020	3	.000
2)	**Root Mean Squared Error (RMSEA)**			.000
	90% Confidence Interval: Lower Bound			.000
	90% Confidence Interval: Upper Bound			.000
	Probability RMSEA <= 0.05			1.000
3)	**Akaike's Information Criterion (AIC)**			61732
	Bayesian Information Criterion (BIC)			61797
4)	**Baseline Comparison**			
	Comparative Fit Index (CFI)			1.000
	Tucker-Lewis Index (TLI)			1.000
5)	**Size of Residuals**			
	Stand. Root Mean Squared Residual (SRMR)			.000
	Coefficient of Determination (CD)			.822

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.810
(Cronbach's alpha = .710)	
McDonald's Omega	.811

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.60
Factor 2	-.12
Factor 3	-.16

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
cult2	0.71	.006	0.70	0.72
cult4	0.83	.006	0.82	0.84
cult5	0.76	.006	0.75	0.77

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
cult2	1.7	0.9	1	4	10920
cult4	1.3	0.6	1	4	10912
cult5	1.6	0.7	1	4	10905

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
cult2	1.78	-0.22	2.34	3.88
cult4	2.67	2.72	5.12	6.48
cult5	2.12	0.22	3.50	5.18

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	161	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	12	4	.020
Strong Invariance (Plus Equal Intercepts)	61	4	.000
Strict Invariance (Plus Equal Error Variances)	36	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	.999
French vs. Italian Language Version	1.000
Italian vs. German Language Version	.999

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.999
Language: Italian	.995

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
culthigh_fs	0.0	0.8	-0.9	2.6	10926

Share of Cases with Imputed Missing Values: 0.3%
(Equivalence of Scores from Robust MLMV: CD = .983)
(Equivalence of Scores from Two-Step Approach: CD = .886)

Scale: Embodied Cultural Capital

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	957	9	.000
	Baseline vs. Saturated	30205	15	.000

2)	**Root Mean Squared Error (RMSEA)**	.098
	90% Confidence Interval: Lower Bound	.093
	90% Confidence Interval: Upper Bound	.103
	Probability RMSEA <= 0.05	.000
3)	**Akaike's Information Criterion (AIC)**	115211
	Bayesian Information Criterion (BIC)	115342
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	.969
	Tucker-Lewis Index (TLI)	.948
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.032
	Coefficient of Determination (CD)	.886

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.872
(Cronbach's alpha = .825)	
McDonald's Omega	.875

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	3.16
Factor 2	.08
Factor 3	-.04
Factor 4	-.04
Factor 5	-.12
Factor 6	-.12

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
manners1	0.54	.007	0.52	0.55
manners2	0.81	.004	0.80	0.82
manners3	0.75	.005	0.74	0.76
verbskill1	0.75	.005	0.74	0.76
verbskill2	0.78	.005	0.77	0.79
verbskill3	0.75	.005	0.74	0.76

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
manners1	3.0	0.8	1	4	10926
manners2	3.1	0.7	1	4	10915
manners3	3.1	0.7	1	4	10915
verbskill1	3.0	0.7	1	4	10927
verbskill2	3.0	0.8	1	4	10922
verbskill3	2.9	0.7	1	4	10901

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
manners1	1.22	-3.60	-1.88	1.24
manners2	2.64	-6.78	-2.90	2.05
manners3	2.15	-6.06	-2.82	1.57
verbskill1	2.14	-5.34	-2.02	1.84
verbskill2	2.39	-5.68	-2.07	1.77
verbskill3	2.16	-5.31	-1.79	2.20

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	548	54	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	32	10	.000
Strong Invariance (Plus Equal Intercepts)	52	10	.000
Strict Invariance (Plus Equal Error Variances)	143	10	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.997
Italian vs. German Language Version	.997

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.999

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
inccap_fs	0.0	0.9	-3.2	1.9	10941

Share of Cases with Imputed Missing Values: 0.7%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .989)

Scale: Embodied Cultural Capital: Manners

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	8814	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	61925
	Bayesian Information Criterion (BIC)	61991
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.801

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.764
(Cronbach's alpha = .685)	
McDonald's Omega	.771

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.41
Factor 2	-.10
Factor 3	-.18

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
manners1	0.59	.008	0.58	0.61
manners2	0.75	.007	0.73	0.76
manners3	0.83	.007	0.82	0.84

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
manners1	3.0	0.8	1	4	10926
manners2	3.1	0.7	1	4	10915
manners3	3.1	0.7	1	4	10915

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
manners1	1.40	-3.78	-1.99	1.33
manners2	2.14	-5.95	-2.57	1.83
manners3	2.86	-7.27	-3.48	1.94

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	325	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	10	4	.046
Strong Invariance (Plus Equal Intercepts)	20	4	.001
Strict Invariance (Plus Equal Error Variances)	32	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.997
Italian vs. German Language Version	.998

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	.998
Language: Italian	.995

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
manners_fs	0.0	0.8	-2.8	1.5	10939

Share of Cases with Imputed Missing Values: 0.5%
(Equivalence of Scores from Robust MLMV: CD = .998)
(Equivalence of Scores from Two-Step Approach: CD = .987)

Scale: Embodied Cultural Capital: Verbal Skills

Monitoring Module

Model and Fit Statistics

		chi2	df	p > chi2
1)	**Likelihood-Ratio Tests**			
	Model vs. Saturated	0	0	
	Baseline vs. Saturated	11399	3	.000

2)	**Root Mean Squared Error (RMSEA)**	.000
	90% Confidence Interval: Lower Bound	.000
	90% Confidence Interval: Upper Bound	.000
	Probability RMSEA <= 0.05	1.000
3)	**Akaike's Information Criterion (AIC)**	62273
	Bayesian Information Criterion (BIC)	62339
4)	**Baseline Comparison**	
	Comparative Fit Index (CFI)	1.000
	Tucker-Lewis Index (TLI)	1.000
5)	**Size of Residuals**	
	Stand. Root Mean Squared Residual (SRMR)	.000
	Coefficient of Determination (CD)	.820

Reliability and Dimensionality

Ordinal Cronbach's Alpha	.817
(Cronbach's alpha = .758)	
McDonald's Omega	.817

Test of (One-)Dimensionality (Parallel Analysis)

Criterion: Retain Factors with Adj. Eigenvalue > 0

	Adjusted Eigenvalue
Factor 1	1.64
Factor 2	-.13
Factor 3	-.16

Standardized Factor Loadings

Indicators	Coef.	(SE)	[95% Conf.	Interval]
verbskill1	0.73	.006	0.72	0.75
verbskill2	0.80	.006	0.79	0.81
verbskill3	0.78	.006	0.77	0.80

Item Descriptives

Indicators	Mean	Std. Dev.	Min.	Max.	Valid Obs.
verbskill1	3.0	0.7	1	4	10927
verbskill2	3.0	0.8	1	4	10922
verbskill3	2.9	0.7	1	4	10901

Parameters of Generalized Structural Equation Model (Ordinal Logit Link)

Indicators	Coef.	Cut1	Cut2	Cut3
verbskill1	2.02	-5.20	-1.96	1.80
verbskill2	2.53	-5.94	-2.15	1.88
verbskill3	2.41	-5.73	-1.93	2.40

Tests and Indices of Factorial Invariance across Survey Languages *

Equality of Variance-Covariance Matrices	chi2	df	p > chi2
	155	18	.000

Tests of Measurement Invariance	chi2	df	p > chi2
Metric Invariance (Equal Factor Loadings)	11	4	.027
Strong Invariance (Plus Equal Intercepts)	25	4	.000
Strict Invariance (Plus Equal Error Variances)	65	4	.000

Configural Factor Similarity

Tucker's Congruence Coefficient	TCC
German vs. French Language Version	1.000
French vs. Italian Language Version	.996
Italian vs. German Language Version	.997

Factor Score Equivalence: Group Specific vs. Invariant Models

Coefficient of Determination	CD
Language: German	1.000
Language: French	1.000
Language: Italian	.992

Factor Score Descriptives

Variable Name	Mean	Std. Dev.	Min.	Max.	Obs.
verbskill_fs	0.0	0.9	-2.7	1.6	10936

Share of Cases with Imputed Missing Values: 0.4%
(Equivalence of Scores from Robust MLMV: CD = .999)
(Equivalence of Scores from Two-Step Approach: CD = .992)

List of Sources

AEQ-M (short-version)

Pekrun, R. , Goetz, T. & Frenzel, A. C. (2005). Achievement emotions questionnaire-mathematics (AEQ-M). User's manual. Department of Psychology, University of Munich.

Böhm-Kasper et al., 2000 (EBI)

Böhm-Kasper, O., Bos, W., Jaeckel, S. & Weishaupt, H. (2000). Skalenhandbuch zur Belastung von Schülern und Lehrern. Das Erfurter Belastungs-Inventar (EBI). Erfurt: Pädagogische Hochschule Erfurt.

COACTIV 2008

Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U., Krauss, S., Kunter, M., Löwen, K., Neubrand, M., & Tsai, Y.-M. (2008). Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV): Dokumentation der Erhebungsinstrumente. Berlin: Max-Planck-Institut für Bildungsforschung.

Eder, 1995

Eder, F. (Eds.) (1995). Das Befinden von Kindern und Jugendlichen in der Schule. Innsbruck: Studienverlag.

GESIS (short-version)

Kovaleva, A., Beierlein, C., Kemper, C. J., & Rammstedt, B. (2012). Eine Kurzskala zur Messung von Kontrollüberzeugung: die Skala Internale-Externale-Kontrollüberzeugung-4 (IE-4) (GESIS-Working Papers, 2012/19). Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften.
Retrieved 12.06.2018 from http://nbn-resolving.de/urn:nbn:de:0168-ssoar-31209

Girnat, 2015 / 2017

Girnat, B. (2017). Gender Differences Concerning Pupils' Beliefs on Teaching Methods and Mathematical Worldviews at Lower Secondary Schools. In: C. Andrà, D. Brunetto, E. Levenson und P. Liljedahl (Eds.):
Teaching and Learning in Maths Classrooms: Emerging Themes in Affect-related Research: Teachers' Beliefs, Students' Engagement and Social Interaction (Research in Mathematics Education). Cham: Springer International Publishing AG, P. 253 – 263.

Girnat, B. (2015). Konstruktivistische und instruktivistische Lehrmethoden aus Schülersicht – Entwicklung eines fragebogenbasierten Erhebungsinstrumentes. In: Beiträge zum Mathematikunterrich, P. 308 – 311.

Girnat, 2018

Girnat, Boris (2018). The PISA Mathematics Self-Efficacy Scale: Questions of Dimensionality and a Latent Class Concerning Algebra. In H. Palmér G J. Skott (Eds), Students' and Teachers' Values, Attitudes, Feelings and Beliefs in Mathematics Classrooms. Springer International Publishing AG.

Hascher, 2004

Hascher, T. (2004). Wohlbefinden in der Schule. Münster: Waxmann.

Hagenauer & Hascher, 2012 (modified)

Hagenauer, G., & Hascher, T. (2012). Erfassung kognitiver Regulationsstrategien bei Schulunlust. Empirische Pädagogik, 26(4), 452-478.

ICILS 2013

Fraillon, J., Ainley, J., Schulz, W., Friedman, T., Gebhardt, E. (2014). Preparing for Life in a Digital Age. The IEA International Information Literacy Study International Report. Cham: Springer.

IGLU 2001

Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (2007): Internationale Grundschul-Lese-Untersuchung 2001 (IGLU/PIRLS 2001). Version: 1. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz. Retrieved 12.06.2018 from http://doi.org/10.5159/IQB_IGLU_2001_v1

Jerusalem & Schwarzer, 1999

Jerusalem, M. & Schwarzer, R. (1999). Allgemeine Selbstwirksamkeitserwartung. In: R. Schwarzer & M. Jerusalem (Eds.). Skalen zur Erfassung von Lehrer und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschamlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Berlin: Freie Universität Berlin.

List of Sources

Köller et al. 2000

Köller, O., Schnabel, K., & Baumert, J. (2000). Der Einfluss der Leistungsstärke von Schulen auf das fachspezifische Selbstkonzept der Begabung und das Interesse. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 32(2), 70-80.

PISA 2012

OECD (2014). PISA 2012 technical report. Paris, Frankreich: OECD.

PISA 2006

OECD (2009). PISA 2006 technical report. Paris, Frankreich: OECD.

PISA 2000

Adams, R. and Wu, M. (2002). PISA 2000 technical report. Paris: OECD.

Rammstedt et al., 2014

Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C. & Kovaleva, A., (2014). Big Five Inventory (BFI-10). Zusammenstellung sozialwissenschaftlicher Items und Skalen. Doi:10.6102/zis76

Rosenberg 1979 (translated)

Rosenberg, M. (1979). Conceiving the self. New York: Basic Books.

Ryan & Conell, 1989

Ryan, R., & Connell, J. (1989). Perceived locus of causality and internalization: Examing reasons for acting in two domains. Journal of Personality and Social Psychology, 57, 749-761.

SELLMO 2012

Spinath, B., Stiensmeier-Pelster, J., Schöne, C., & Dickhäuser, O. (2002). Skalen zur Erfassung der Lern- und Leistungsmotivation (SELLMO). Göttingen: Hogrefe.

Seidel, Prenzel & Kobarg, 2005

Seidel, T., Prenzel, M., & Kobarg, M. (Eds.). (2005). How to run a video study: Technical report of the IPN Video Study. Münster: Waxmann.

Sen, 1985

Sen, A.K. (1985). Commodities and Capabilities, Oxford: Elsevier

Szydlik, 2008

Szydlik, M. (2008). Intergenerational Solidarity and Conflict. Journal of Comparative Family Studies, 39(1), 97-114.

TREE1

TREE (2016). Concepts and Scales. Survey waves 1 to 9, 2001-2014. Berne: TREE.

TREE2

Hupka-Brunner, S., Jann, B., Meyer, T., Imdorf, C., Sacchi, S., Müller, B., Scharenberg, K., von Rotz, C., Koomen, M. & Becker, R. (2015). Erläuterungen zum Kontextfragebogen der ÜGK 2016: Allgemeiner Teil. Bern.

Watermann, 2000

Watermann, R. (2000). Berufliche Wertorientierungen im Wandel. Eine Kohortenanalyse zur Dynamik arbeitsbezogener Einstellungen anhand von ALLBUS-Umfragedaten. Münster: Institut für sozialwissenschaftliche Forschung e.V